

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:45 ; Search time 31.95 Seconds
(without alignments)
58.041 Million cell updates/sec

Title: US-09-851-422B-1
Perfect score: 46
Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	456	4 Q14316	Q14316 homo sapien
2	46	100.0	461	6 Q95ND7	Q95ND7 pan troglod
3	46	100.0	461	6 Q95ND6	Q95ND6 pan troglod
4	36	78.3	908	5 Q96516	Q96516 caenorhabdi
5	36	78.3	909	5 Q76618	Q76618 caenorhabdi
6	34	73.9	60	5 Q9U4L5	Q9U4L5 tetrahymena
7	34	73.9	254	3 Q01136	Q01136 metarhizium
8	34	73.9	255	3 Q9Y7A9	Q9Y7A9 metarhizium
9	34	73.9	256	3 Q9Y842	Q9Y842 metarhizium
10	34	73.9	434	16 Q8UK17	Q8UK17 agrobacteri
11	33	71.7	70	2 Q45148	Q45148 bacteroides
12	33	71.7	165	16 Q9I6M8	Q9I6M8 pseudomonas
13	33	71.7	212	10 Q8W599	Q8W599 secale cere
14	33	71.7	212	10 Q8W236	Q8W236 triticum ae
15	33	71.7	252	2 Q8VW87	Q8VW87 terrabacter
16	33	71.7	253	4 Q8WZB4	Q8WZB4 homo sapien

ALIGNMENTS

RESULT 1

Q14316 ID Q14316 PRELIMINARY; PRT; 456 AA.

AC Q14316; DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-AUG-1999 (Tremblrel. 11, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE F9 (Coagulation factor IX (Plasma THROMBOPLASTIC component, Christmas disease, HAEMOPHILIA B)) (Factor IX).
GN F9 OR FACTOR IX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 3-19 FROM N.A.
RX MEDLINE=88327116; PubMed=3416069;
RA Reitsma P.A., Bertina R.M., Ploos van Amstel J.K., Riemsma A., Briet E.;
RT "The putative factor IX gene promoter in hemophilia B Leyden.";
RL Blood 72:1074-1076(1988).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; AL033403; CAA21954.1; -;
DR EMBL; X55008; CAB38245.2; -;
DR HSSP; P00740; 1CPH.
DR MEROPS; S01.214; -;
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001254; Ser_protease_Try.

Q9W7P9 paralichthy
Q9VXR2 drosophila
Q9V660 drosophila
Q42723 emericella
Q9WU50 mus musculus
Q9GV23 sarcophaga
Q9IZU8 mus musculus
Q9FTY5 oryza sativ
Q8TUD2 methanosarc
Q8TUD1 methanosarc
Q97F76 clostridium
Q91NR8 arabidopsis
Q97C14 thermoplaam
Q8YC01 brucella me
Q9ZW73 arabidopsis
Q8Y32 arabidopsis
Q9DGT6 turkey herp
Q9UR58 rhizomucor
Q91GQ3 oryza sativ
Q55323 serratia sp
Q98139 escherichia
Q54478 serratia ma
Q9VFE3 aeropyrum p
Q9RV11 deinococcus
Q9W3W8 drosophila
Q961J3 homo sapien
Q9F2B7 thauera aro
Q9AE05 glycine max
Q9HF62 ashbya goss

```
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00001; GLABLOOD.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM0069; GLA; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW Serine protease.
SQ SEQUENCE 456 AA; 51149 MW; 54E20A1B3964E234 CRC64;

Query Match 100.0%; Score 46; DB 4; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 371 LVDRATCLR 379

RESULT 2
Q95ND7 PRELIMINARY; PRT; 461 AA.
AC Q95ND7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Coagulation factor XI.
GN F9.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=504;
RA Satta Y.;
RT "Comparison of DNA and protein polymorphisms between humans and
RT chimpanzees.";
RL Genes Genet. Syst. 0:0-0(2001).
DR EMBL; AB062470; BAB58885.1; JOINED.
DR EMBL; AB062458; BAB58885.1; JOINED.
DR EMBL; AB062460; BAB58885.1; JOINED.
DR EMBL; AB062462; BAB58885.1; JOINED.
DR EMBL; AB062464; BAB58885.1; JOINED.
DR EMBL; AB062466; BAB58885.1; JOINED.
DR EMBL; AB062468; BAB58885.1; JOINED.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 461 AA; 51695 MW; 8F5A69A525DF65B5 CRC64;

Query Match 100.0%; Score 46; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 4
Q96516 PRELIMINARY; PRT; 908 AA.
ID Q96516
```

```
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 461 AA; 51764 MW; 30C2F857C0F77F45 CRC64;

Query Match 100.0%; Score 46; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 3
Q95ND6 PRELIMINARY; PRT; 461 AA.
AC Q95ND6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Coagulation factor XI.
GN F9.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=505;
RA Satta Y.;
RT "Comparison of DNA and protein polymorphisms between humans and
RT chimpanzees.";
RL Genes Genet. Syst. 0:0-0(2001).
DR EMBL; AB062471; BAB58886.1; JOINED.
DR EMBL; AB062459; BAB58886.1; JOINED.
DR EMBL; AB062461; BAB58886.1; JOINED.
DR EMBL; AB062463; BAB58886.1; JOINED.
DR EMBL; AB062465; BAB58886.1; JOINED.
DR EMBL; AB062467; BAB58886.1; JOINED.
DR EMBL; AB062469; BAB58886.1; JOINED.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 461 AA; 51695 MW; 8F5A69A525DF65B5 CRC64;

Query Match 100.0%; Score 46; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 4
Q96516 PRELIMINARY; PRT; 908 AA.
ID Q96516
```


OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarhizium.
OX NCBI_TaxID=5530;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=MEL;
RX MEDLINE=96105219; PubMed=8529882;
RA Smithson S.L., Paterson I.C., Bailey A.M., Screen S.E., Hunt B.A.,
Cobb B.D., Cooper R.M., Charnley A.K., Clarkson J.M.;
RT "Cloning and characterization of a gene encoding a cuticle-degrading
protease from the insect pathogenic fungus *Metarhizium anisopliae*.";
RL Gene 166:161-165(1995).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; X78875; CAA55477.1; -.
DR HSSP; P35049; 1TRY.
DR MEROPS; S01.103; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 254 AA; 26117 MW; 0AED96A5C52DBA54 CRC64;
Query Match 73.9%; Score 34; DB 3; Length 254;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATC 7
Db 169 VVDRATC 175
:|||||
RESULT 8
QY7A9
ID QY7A9 PRELIMINARY; PRT; 255 AA.
AC QY7A9;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Trypsin-related protease.
GN TRY2.
OS Metarhizium anisopliae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarhizium.
OX NCBI_TaxID=5530;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=MEL;
RX Screen S.E., St Leger R.J.;
RT "Isolation of multiple protease genes from the entomopathogenic fungus
Metarhizium anisopliae.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; AF130865; RAD29675.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.103; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 255 AA; 26289 MW; 88D979ED300B4B7 CRC64;

Query Match 73.9%; Score 34; DB 3; Length 255;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATC 7
Db 173 VVDRATC 179
:|||||
RESULT 9
QY842
ID QY842 PRELIMINARY; PRT; 256 AA.
AC QY842;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Trypsin-related protease precursor.
GN TRY1.
OS Metarhizium anisopliae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarhizium.
OX NCBI_TaxID=5530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEL;
RX Screen S.E., St Leger R.J.;
RT "Isolation of multiple protease genes from the entomopathogenic fungus
Metarhizium anisopliae.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; AJ242736; CAB44652.1; -.
DR HSSP; P35049; 1TRY.
DR MEROPS; S01.103; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 30 256 TRYPSIN-RELATED PROTEASE.
SQ SEQUENCE 256 AA; 26201 MW; 34696608745CB982 CRC64;
Query Match 73.9%; Score 34; DB 3; Length 256;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATC 7
Db 171 VVDRATC 177
:|||||
RESULT 10
QY817
ID QY817 PRELIMINARY; PRT; 434 AA.
AC QY817;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Citrate synthase.
GN CIS2 OR ATU5307 OR AGR PAT 441.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG Plasmid AT.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Doan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestor E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58";
 RL Science 294:2317-2323(2001).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmiel K., Gordon J., Vaundin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 DR EMBL; A5008952; AAL45995.1; --
 DR EMBL; A5007899; AAK90682.1; ALT_INIT.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 434 AA; 48002 MW; C5FDBDFA56B01E28 CRC64;

Query Match. 73.9%; Score 34; DB 16; Length 434;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATCL 9
 Db 14 LIQRETCL 22

RESULT 11
 Q45148 PRELIMINARY; PRT; 70 AA.
 ID Q45148
 AC Q45148
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE ORF1169b.
 OS Bacteroides fragilis.
 OC Bacteria; CF8 group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
 OC Bacteroides.
 OC NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95291442; PubMed=7773395;
 RA Trinh S., Haggoud A., Reyssat G., Sebald M.;
 RT "Plasmids pIP419 and pIP421 from Bacteroides: 5-nitroimidazole
 RT resistance genes and their upstream insertion sequence elements";
 RL Microbiology 141:927-935(1995).
 DR EMBL; X76949; CAA54271.1; --
 SQ SEQUENCE 70 AA; 8173 MW; 2C8A79953A919FF3 CRC64;

Query Match 71.7%; Score 33; DB 2; Length 70;
 Best Local Similarity 62.5%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCL 8
 Db 57 MLQKATCL 64

RESULT 12
 Q916M8 PRELIMINARY; PRT; 165 AA.
 ID Q916M8

AC Q916M8;
 DT 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMELrel. 18, Last annotation update)
 DE Hypothetical protein PA0261.
 GN PA0261.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004464; AAG03650.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 18391 MW; C93ED41A0E251C7B CRC64;

Query Match 71.7%; Score 33; DB 16; Length 165;
 Best Local Similarity 75.0%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCL 8
 Db 145 LLDEATCL 152

RESULT 13
 Q8W599 PRELIMINARY; PRT; 212 AA.
 ID Q8W599
 AC Q8W599
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE CBF-like protein.
 DE CBF-like protein.
 OS Secale cereale (Rye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Secale.
 OC NCBI_TaxID=4550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21563065; PubMed=11706173;
 RA Jaglo K.R., Kleff S., Amundsen K.L., Zhang X., Haake V., Zhang J.Z.,
 RA Deits T., Thomashow M.F.;
 RT "Components of the Arabidopsis C-Repeat/Dehydration-Responsive Element
 RT Binding Factor Cold-Response Pathway Are Conserved in Brassica napus
 RT and Other Plant Species";
 RL Plant Physiol. 127:910-917(2001).
 DR EMBL; AF370728; AAL35759.1; --
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR ProDom; PD001423; TF_AP2; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 212 AA; 23291 MW; E1D0265030102159 CRC64;

Query Match 71.7%; Score 33; DB 10; Length 212;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCL 8
 Db 86 LLDRATCL 93

```
RESULT 14
Q8W236
ID Q8W236 PRELIMINARY; PRT; 212 AA.
AC Q8W236;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative CRT/DRE-binding factor.
GN CBF1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
SEQUENCE FROM N.A.
RP
RC STRAIN=CV. WINOKA;
RX MEDLINE=21563065; PubMed=11706173;
RA Jaglo KR., Kleff S., Amundsen K.L., Zhang X., Haake V., Zhang J.Z.,
RA Deits T., Thomasow M.F.;
RT "Components of the Arabidopsis C-Repeat/Dehydration-Responsive Element
RT Binding Factor Cold-Response Pathway Are Conserved in Brassica napus
RT and Other Plant Species.";
RL Plant Physiol. 127:910-917(2001).
DR EMBL; AF376136; AAL37944.1; -.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRODOM; PD001423; TF_AP2; 1.
DR SMART; SM00380; AP2; 1.
DR SQ SEQUENCE 212 AA; 23336 MW; 585C2E6DD86EA28D CRC64;

Query Match 71.7%; Score 33; DB 10; Length 212;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRAATCL 8
Db 86 LLDBAACL 93

RESULT 15
Q8VV87
ID Q8VV87 PRELIMINARY; PRT; 252 AA.
AC Q8VV87;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LexA repressor-like protein.
GN ORFL6.
OS Terrabacter sp.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Intrasporangiaceae; Terrabacter.
OX NCBI_TaxID=54015;
RN [1]
SEQUENCE FROM N.A.
RP
RC STRAIN=DF63;
RA Kasuga K., Nojiri H., Yanane H., Kodama T., Omori T.;
RT "Cloning and characterization of genes involved in the degradation of
RT dibenzofuran by Terrabacter sp. strain DBF63.";
RL J. Ferment. Bioeng. 84:387-399(1997).
DR EMBL; AB004563; BAB7870.1; -.
DR InterPro; IPR001647; HTH_Tetr.
DR InterPro; IPR000129; Peptidase_S24.
DR Pfam; PF00717; Peptidase_S24; 1.
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PRO0455; HTHETR.
DR PRINTS; PRO0726; LEXASERPTASE.
DR SQ SEQUENCE 252 AA; 27167 MW; B7C5BACE598ADBA3 CRC64;

Query Match 71.7%; Score 33; DB 2; Length 252;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LVDRAATC 7
Db 225 LVDRAATC 231

Search completed: May 6, 2003, 15:02:46
Job time : 34.95 secs
```

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:24 ; Search time 9 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46

Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	271	1 FA9_PIG	P16293 sus scrofa
2	46	100.0	274	1 FA9_SHEEP	P16291 ovis aries
3	46	100.0	282	1 FA9_RAT	P16296 rattus norv
4	46	100.0	285	1 FA9_CAVPO	P16295 cavia porce
5	46	100.0	416	1 FA9_BOVIN	P00741 bos taurus
6	46	100.0	452	1 FA9_CANFA	P19540 canis famil
7	46	100.0	459	1 FA9_MOUSE	P16294 mus musculu
8	46	100.0	461	1 FA9_HUMAN	P00740 homo sapien
9	42	91.3	275	1 FA9_RABIT	P16292 oryctolagus
10	36	78.3	586	1 VP40_ILTVT	P23984 infectious
11	34	73.9	2347	1 KR05_HUMAN	P08922 homo sapien
12	33	71.7	123	1 UL59_HGWA	P16746 human cytom
13	33	71.7	253	1 CFAD_HUMAN	P00746 homo sapien
14	33	71.7	546	1 Y220_HUMAN	Q92817 homo sapien
15	32	69.6	268	1 CLCR_HUMAN	Q99895 homo sapien
16	32	69.6	311	1 L59_MICGR	P43433 micromonosp
17	32	69.6	323	1 CMGB_BACSU	P25954 bacillus su
18	32	69.6	547	1 ICA3_HUMAN	P32942 homo sapien
19	32	69.6	688	1 CHLB_CHLRE	P36437 chlamydomon
20	32	69.6	863	1 LDVR_CHICK	P98165 gallus gall
21	32	69.6	2704	1 BP41_HUMAN	Q03001 homo sapien
22	31	67.4	234	1 MAUM_METFL	Q50423 methylobaci
23	31	67.4	378	1 PROB_BRUME	Q8Y179 brucella me
24	31	67.4	431	1 SC65_RAT	Q64375 rattus norv
25	31	67.4	437	1 NO55_HUMAN	Q92791 homo sapien
26	31	67.4	873	1 LDVR_HUMAN	P98155 homo sapien
27	31	67.4	873	1 LDVR_MOUSE	P98156 mus musculu
28	31	67.4	873	1 LDVR_RABIT	P35953 oryctolagus
29	31	67.4	873	1 LDVR_RAT	P98166 rattus norv
30	30	65.2	89	1 RT15_BOVIN	P82913 bos taurus
31	30	65.2	103	1 CHLB_LYCAN	P37847 lycopodium
32	30	65.2	103	1 CHLB_LYCCO	P37848 lycopodium
33	30	65.2	212	1 TBAA_PNECA	Q07972 pneumocysti

RESULT 1

ID	FA9_PIG	STANDARD	PRT	271 AA
AC	P16293			
DT	01-AUG-1990	(Rel. 15, Created)		
DT	01-AUG-1990	(Rel. 15, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Fragment).			
GN	F9.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
OX	NCBI_TaxID=9823			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90152675; PubMed=2303254;			
RA	Sarkar G., Koerber D.D., Sommer S.S.;			
RT	"Direct sequencing of the activation peptide and the catalytic domain			
RT	of the factor IX gene in six species."			
RL	Genomics 6:133-143(1990).			
RN	[2]			
RP	SEQUENCE OF 98-106 FROM N.A.			
RC	STRAIN=Meishan, and wild boar;			
RX	MEDLINE=97009812; PubMed=8856916;			
RA	Singer E.N., Armour J.A.L., Jeffreys A.J.;			
RT	"Detection of an Mb01 RFLP at the porcine clotting factor IX locus			
RT	and verification of sex linkage."			
RL	Anim. Genet. 27:130-130(1996).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE=96003866; PubMed=7568220;			
RA	Brandstetter H., Bauer M., Huber R., Lollar P., Bode W.;			
RT	"X-ray structure of clotting factor IXa: active site and module			
RT	structure related to Xase activity and hemophilia B."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9796-9800(1995).			
CC	FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT			
CC	PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY			
CC	CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++			
CC	IONS, PHOSPHOLIPIDS, AND FACTOR VIII.			
CC	FUNCTIONAL ACTIVITY: Hydrolyzes one Arg-Ile bond in factor X to			
CC	form factor Xa.			
CC	SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE			
CC	ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2			
CC	CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.			
CC	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
DR	EMBL; M26235; AAA31031.1; -			
DR	EMBL; X92427; CAA63155.1; -			

ALIGNMENTS

34	30	65.2	305	1	S17C_RAT	Q64686	rattus norv
35	30	65.2	346	1	Y161_XANCV	P25438	xanthomonas
36	30	65.2	430	1	CINA_MYCTU	O07731	mycobacteri
37	30	65.2	444	1	CHLB_CHLPT	P37824	chlamydomon
38	30	65.2	445	1	TBA3_YEAST	P09734	saccharomyc
39	30	65.2	447	1	TBA1_YEAST	P09733	saccharomyc
40	30	65.2	449	1	TBA2_SCHPO	P04689	schizosacch
41	30	65.2	451	1	TBA2_EMENI	P24634	emericeilla
42	30	65.2	475	1	FA10_CHICK	P25155	gallus gall
43	30	65.2	482	1	RNF9_HUMAN	Q9udy6	homo sapien
44	30	65.2	492	1	FA10_BOVIN	P00743	bos taurus
45	30	65.2	546	1	PUT2_AGABI	P78568	agaricus bi

```
DR EMBL; X02593; CAA63337.1; -.
DR PDB; 1PFX; 17-AUG-96.
DR MEROPS; S01.214; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; PARTIAL.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
KW Hydrolase; Glycoprotein; 3D-structure.
FT NON_TER 1 1
FT ACT_SITE 85 85 CHARGE RELAY SYSTEM.
FT ACT_SITE 133 133 CHARGE RELAY SYSTEM.
FT ACT_SITE 229 229 CHARGE RELAY SYSTEM.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 271 271
SQ SEQUENCE 271 AA; 23992 MW; 3D6C18P7FC66A24B CRC64;

Query Match 100.0%; Score 46; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 194 LVDRATCLR 202

RESULT 2
FA9_SHEEP STANDARD; PRT; 274 AA.
AC P16291;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor IX (EC 3.4.21.22), (Christmas factor) (Fragment).
GN F9.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90152675; PubMed=2303254;
RA Sarkar G., Koeberl D.D., Sommer S.S.;
RT "Direct sequencing of the activation peptide and the catalytic domain
of the factor IX gene in six species.";
RL Genomics 6:133-143(1990).
CC -!- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
CC PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
CC CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
CC IONS, PHOSPHOLIPIDS, AND FACTOR VIII.
CC -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC form factor Xa.
CC -!- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE
CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M26233; AAA31520.1; -.
DR EMBL; M26233; 1PFX.
```

```
DR HSSP; P16293; 1PFX.
DR MEROPS; S01.214; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; PARTIAL.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
KW Hydrolase; Glycoprotein.
FT NON_TER 1 1
FT ACT_SITE 89 89 CHARGE RELAY SYSTEM.
FT ACT_SITE 137 137 CHARGE RELAY SYSTEM.
FT ACT_SITE 233 233 CHARGE RELAY SYSTEM.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 274 274
SQ SEQUENCE 274 AA; 30595 MW; D3617FC3B1D33E9B CRC64;

Query Match 100.0%; Score 46; DB 1; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 198 LVDRATCLR 206

RESULT 3
FA9_RAT STANDARD; PRT; 282 AA.
AC P16296;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor IX (EC 3.4.21.22), (Christmas factor) (Fragment).
GN F9 OR CF9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90152675; PubMed=2303254;
RA Sarkar G., Koeberl D.D., Sommer S.S.;
RT "Direct sequencing of the activation peptide and the catalytic domain
of the factor IX gene in six species.";
RL Genomics 6:133-143(1990).
CC -!- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
CC PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
CC CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
CC IONS, PHOSPHOLIPIDS, AND FACTOR VIII.
CC -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC form factor Xa.
CC -!- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE
CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M26247; AAA41162.1; -.
DR HSSP; P16293; 1PFX.
```


DR MEROPS; S01.214; --
 DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00011; GLU CARBOXYLATION; PARTIAL.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
 KW Hydrolase; Glycoprotein.
 FT NON_TER 1 1
 FT ACT_SITE 96 96 CHARGE RELAY SYSTEM.
 FT ACT_SITE 144 144 CHARGE RELAY SYSTEM.
 FT ACT_SITE 240 240 CHARGE RELAY SYSTEM.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 282 282
 SQ SEQUENCE 282 AA; 31447 MW; 88B37B0A4673BEC9 CRC64;
 Query Match 100.0%; Score 46; DB 1; Length 282;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 |||||
 DB 205 LVDRATCLR 213
 RESULT 4
 FA9_CAVPO
 ID FA9_CAVPO STANDARD; PRT; 285 AA.
 AC P16295;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Fragment).
 GN F9.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90152675; PubMed=2301254;
 RA Sarkar G., Koerber D.D., Sommer S.S.;
 RT "Direct sequencing of the activation peptide and the catalytic domain
 of the factor IX gene in six species.";
 RL Genomics 6:133-143(1990).
 CC -1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
 PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
 CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
 IONS, PHOSPHOLIPIDS, AND FACTOR VIIIa.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg--Ile bond in factor X to
 form factor Xa.
 CC -1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIa, WHICH EXCISES THE
 ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
 CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC ENBL; M26237; AAA37037.1; --
 DR HSP3; P16293; 1PFX.
 DR MEROPS; S01.214; --

DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00011; GLU CARBOXYLATION; PARTIAL.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
 KW Hydrolase; Glycoprotein.
 FT NON_TER 1 1
 FT ACT_SITE 99 99 CHARGE RELAY SYSTEM.
 FT ACT_SITE 147 147 CHARGE RELAY SYSTEM.
 FT ACT_SITE 243 243 CHARGE RELAY SYSTEM.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 285 285
 SQ SEQUENCE 285 AA; 32032 MW; 68B36DD317C11C60 CRC64;
 Query Match 100.0%; Score 46; DB 1; Length 285;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 |||||
 DB 208 LVDRATCLR 216
 RESULT 5
 FA9_BOVIN
 ID FA9_BOVIN STANDARD; PRT; 416 AA.
 AC P00741;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Coagulation factor IX (EC 3.4.21.22) (Christmas factor).
 GN F9.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=80056619; PubMed=291916;
 RA Katayama K., Ericsson L.H., Enfield D.L., Walsh K.A., Neurath H.,
 Davie E.W., Titani K.;
 RT "Comparison of amino acid sequence of bovine coagulation Factor IX
 (Christmas Factor) with that of other vitamin K-dependent plasma
 proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:4990-4994(1979).
 RN [2]
 RP REVISION TO 64.
 RX MEDLINE=83308813; PubMed=6688526;
 RA McMullen B.A., Fujikawa K., Kisiel W.;
 RT "The occurrence of beta-hydroxyaspartic acid in the vitamin
 K-dependent blood coagulation zymogens.";
 RL Biochem. Biophys. Res. Commun. 115:8-14(1983).
 RN [3]
 RP SEQUENCE OF 51-111 FROM N.A.
 RX MEDLINE=8272386; PubMed=6287289;
 RA Choo K.H., Gould K.G., Rees D.J.G., Brownlee G.G.;
 RT "Molecular cloning of the gene for human anti-haemophilic factor IX.";
 RL Nature.299:178-180(1982).
 RN [4]
 RP STRUCTURE OF CARBOHYDRATE ON SER-53.
 RX MEDLINE=89213999; PubMed=3149637;
 RA Hase S., Kawabata S., Nishimura H., Takeya H., Sueyoshi T.,
 Miyata T., Iwanaga S., Takao T., Shimoniishi Y., Ikenaka T.;
 RT "A new trisaccharide sugar chain linked to a serine residue in bovine
 blood coagulation factors VII and IX.";

RC TISSUE=Liver;
RX MEDLINE=89323338; PubMed=2752110;
RA Evans J.P., Watzke H.H., Ware J.L., Stafford D.W., High K.A.;
RT "Molecular cloning of a cDNA encoding canine factor IX.";
RL Blood 74:207-212(1989).
RN [3]
RP VARIANT HEMOPHILIA B GLU-418.
RX MEDLINE=90099303; PubMed=2481310;
RA Evans J.P., Brinkhous K.M., Brayer G.D., Reiser H.M., High K.A.;
RT "Canine hemophilia B resulting from a point mutation with unusual
consequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:10095-10099(1989).
CC -!- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
IONS, PHOSPHOLIPIDS, AND FACTOR VIIIa.
CC -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
form factor Xa.
CC -!- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE
ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
CC -!- DISEASE: HEMOPHILIA B IS DUE TO DEFECTS IN FACTOR IX.
CC -!- MISCELLANEOUS: CALCIUM BINDS TO THE GAMMA-CARBOXYGLUTAMIC ACID
(GLA) RESIDUES AND, WITH STRONGER AFFINITY, TO ANOTHER SITE,
BEYOND THE GLA DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; M21757; AAA75006.1; -;
DR EMBL; M33826; AAA30844.1; -;
DR PIR; A30351; A30351.
DR HSSP; P00740; 1CFH.
DR MEROPS; S01.214; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
KW Hydroxylase; Glycoprotein; Vitamin K; Gamma-carboxyglutamic acid;
KW Hemophilia; Hydroxylation; Zymogen; Signal; EGF-like domain;
KW Repeat; Disease mutation.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 39 FACTOR IXA (ACTIVE FORM) LIGHT CHAIN.
FT CHAIN 40 183
FT PROPEP 184 217 ACTIVATION PEPTIDE.

FT CHAIN 218 452
FT DOMAIN 86 122
FT DOMAIN 123 164
FT DOMAIN 183 218
FT SITE 183 218
FT SITE 217 46
FT MOD_RES 46 46
FT MOD_RES 47 47
FT MOD_RES 47 47
FT MOD_RES 54 54
FT MOD_RES 56 56
FT MOD_RES 59 59
FT MOD_RES 60 60
FT MOD_RES 65 65
FT MOD_RES 66 66
FT MOD_RES 69 69
FT MOD_RES 72 72
FT MOD_RES 75 75
FT MOD_RES 79 79
FT MOD_RES 103 103
FT DISULFID 57 62
FT DISULFID 90 101
FT DISULFID 95 110
FT DISULFID 112 121
FT DISULFID 127 138
FT DISULFID 134 148
FT DISULFID 150 163
FT CARBOHYD 197 197
FT CARBOHYD 207 207
FT CARBOHYD 297 297
FT ACT_SITE 258 258
FT ACT_SITE 306 306
FT ACT_SITE 402 402
FT VARIANT 418 418
SQ SEQUENCE 452 AA; 50827 MW; 1F6537C46A6960ED CRC64;
Query Match 100.0%; Score 46; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
DB 367 LVDRATCLR 375
RESULT 7
FA9 MOUSE STANDARD; PRT; 459 AA.
ID FA9_MOUSE
AC P16294;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor)
DE (fragment).
DE F9 OR CF9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90215309; PubMed=2323576;

RA Wu S.-M., Stafford D.W., Ware J.;
 RT "Deduced amino acid sequence of mouse blood-coagulation factor IX.";
 RN Gene 86:275-278(1990).
 RX SEQUENCE OF 168-451 FROM N.A.
 RA MEDLINE=90152675; PubMed=2303254;
 RA Sarkar G., Koerber D.D., Sommer S.S.;
 RT "Direct sequencing of the activation peptide and the catalytic domain
 of the factor IX gene in six species.";
 RN Genomics 6:133-143(1990).
 CC -1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
 PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
 CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
 IONS, PHOSPHOLIPIDS, AND FACTOR VIII.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-Ile bond in factor X to
 form factor Xa.
 CC -1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE
 ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
 CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M23109; AAA37629.1; -;
 DR EMBL; M26236; AAA37630.1; -;
 DR PIR; JQ0419; JQ0419.
 DR HSSP; P00740; ICFH.
 DR MEROPS; S01.214; -;
 DR MGD; MGI:86384; F9.
 DR InterPro; IPR000152; Aex hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001254; Ser protease_Try.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR Pfam; PF00594; Gla; 1.
 DR SMART; SM00179; EGF_Ca; 1.
 DR SMART; SM00069; GLA; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00010; ASX HYDROXYL; 1.
 DR PROSITE; PS00022; EGF 1; 1.
 DR PROSITE; PS01186; EGF 2; 2.
 DR PROSITE; PS01187; EGF_Ca; 1.
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
 DR PROSITE; PS00240; TRYPSIN DOM; 1.
 DR PROSITE; PS00134; TRYPSIN HIS; 1.
 DR PROSITE; PS00135; TRYPSIN SER; 1.
 KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
 KW Hydroxylase; Glycoprotein; Vitamin K; Gamma-carboxyglutamic acid;
 KW Hemophilia; Hydroxylation; Zymogen; Signal; EGF-like domain;
 KW Repeat.
 FT NON_TER 1 1
 FT SIGNAL <1 16 BY SIMILARITY.
 FT PROPEP 17 34
 FT CHAIN 35 180 FACTOR IXA (ACTIVE FORM) LIGHT CHAIN.
 FT PROPEP 181 224 ACTIVATION PEPTIDE.
 FT CHAIN 225 459 FACTOR IXA (ACTIVE FORM) HEAVY CHAIN.
 FT DOMAIN 81 117 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 118 159 EGF-LIKE 2.
 FT DOMAIN 225 459 SERINE PROTEASE.
 FT SITE 180 181 CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).
 FT SITE 224 225 CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).
 FT MOD_RES 41 41 GAMMA-CARBOXYGLUTAMIC ACID
 (BY SIMILARITY).

FT	MOD_RES	42	42	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	49	49	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	51	51	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	54	54	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	55	55	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	61	61	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	64	64	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	67	67	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	70	70	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	74	74	GAMMA-CARBOXYGLUTAMIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	98	98	HYDROXYLATION (BY SIMILARITY).
FT	ACT_SITE	265	265	CHARGE RELAY SYSTEM.
FT	ACT_SITE	313	313	CHARGE RELAY SYSTEM.
FT	ACT_SITE	409	409	CHARGE RELAY SYSTEM.
FT	DISULFID	52	57	BY SIMILARITY.
FT	DISULFID	85	96	BY SIMILARITY.
FT	DISULFID	90	105	BY SIMILARITY.
FT	DISULFID	107	116	BY SIMILARITY.
FT	DISULFID	122	133	BY SIMILARITY.
FT	DISULFID	129	143	BY SIMILARITY.
FT	DISULFID	145	158	BY SIMILARITY.
FT	CARBOHYD	192	192	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	304	304	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	363	363	H -> Q (IN REF. 2).
FT	CONFLICT	388	388	T -> I (IN REF. 2).
FT	SEQUENCE	459 AA; 51635 MW; EF439C840D6C8C1A		CRC64;

Query Match 100.0%; Score 46; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 0.055;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LVDRATCLR	9
DB	374	LVDRATCLR	382

RESULT 8
 ID_FA9_HUMAN STANDARD; PRT; 461 AA.
 AC P00740;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor).
 GN F9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86000558; PubMed=2994716;
 RX Yoshitake S., Schach B.G., Foster D.C., Davie E.W., Kurachi K.;
 RT "Nucleotide sequence of the gene for human factor IX (antihemophilic
 factor B)";
 RL Biochemistry 24:3736-3750(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190593; PubMed=3857619;
 RA McGraw R.A., Davis L.M., Noyes C.M., Lundblad R.L., Roberts H.R.,

RA Graham J.B., Stafford D.W.;
 RT "Evidence for a prevalent dimorphism in the activation peptide of
 RT human coagulation factor IX.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2847-2851(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84236100; PubMed=6329734;
 RA Anson D.S., Choo K.H., Rees D.J.G., Giannelli F., Gould K.G.,
 RA Huddleston J.A., Brownlee G.G.;
 RT "The gene structure of human anti-haemophilic factor IX.";
 RL EMBO J. 3:1053-1060(1984).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83220788; PubMed=6687940;
 RA Jaye M., de la Salle H., Schamber F., Balland A., Kohli V.,
 RA Findeli A., Tolstoshev P., Lecocq J.P.;
 RT "Isolation of a human anti-haemophilic factor IX cDNA clone using a
 RT unique 52-base synthetic oligonucleotide probe deduced from the amino
 RT acid sequence of bovine factor IX.";
 RL Nucleic Acids Res. 11:2325-2335(1983).
 RN [5]
 RP SEQUENCE OF 36-326 FROM N.A.
 RX MEDLINE=83200526; PubMed=6089357;
 RA Jagadeeswaran P., Lavelle D.E., Kaul R., Mohandas T., Warren S.T.;
 RT "Isolation and characterization of human factor IX cDNA:
 RT identification of Tag I polymorphism and regional assignment.";
 RL Somat. Cell Mol. Genet. 10:465-473(1984).
 RN [6]
 RP SEQUENCE OF 290-359 FROM N.A.
 RX MEDLINE=88127096; PubMed=3340835;
 RA Scofield E.S., Koerber D.D., Sarkar G., Sommer S.S.;
 RT "Genomic amplification with transcript sequencing.";
 RL Science 239:491-494(1988).
 RN [7]
 RP SEQUENCE OF 444-461 FROM N.A.
 RX MEDLINE=94054330; PubMed=8236150;
 RA de la Salle C., Charmanter J.L., Baas M.J., Schwartz A.,
 RA Wiesel M.L., Grunbaum L., Cazenave J.-P.;
 RT "A deletion located in the 3' non translated part of the factor IX
 RT gene responsible for mild haemophilia B.";
 RL Thromb. Haemost. 70:370-371(1993).
 RN [8]
 RP SEQUENCE OF 47-461 (VARIANT NAGOYA).
 RX MEDLINE=90078229; PubMed=2592373;
 RA Suehiro K., Kawabata S.-I., Miyata T., Takeya H., Takamatsu J.,
 RA Ogata K., Kamiya T., Saito H., Niho Y., Iwanaga S.;
 RT "Blood clotting factor IX BM Nagoya. Substitution of arginine 180 by
 RT tryptophan and its activation by alpha-chymotrypsin and rat mast cell
 RT chymase.";
 RL J. Biol. Chem. 264:21257-21265(1989).
 RN [9]
 RP HYDROXYLATION OF ASP-110.
 RX MEDLINE=93308833; PubMed=6689526;
 RA McMullen B.A., Fujikawa K., Kisiel W.;
 RT "The occurrence of beta-hydroxyaspartic acid in the vitamin
 RT K-dependent blood coagulation zymogens.";
 RL Biochem. Biophys. Res. Commun. 115:8-14(1983).
 RN [10]
 RP PROCRESSING, ACTIVE SITE, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=78194509; PubMed=659613;
 RA di Scipio R.G., Kurachi K., Davie E.W.;
 RT "Activation of human factor IX (Christmas factor).";
 RL J. Clin. Invest. 61:1528-1538(1978).
 RN [11]
 RP CALCIUM-BINDING DATA.
 RX MEDLINE=84185715; PubMed=6425296;
 RA Morita T., Isaacs B.S., Esmon C.T., Johnson A.E.;
 RT "Derivatives of blood coagulation factor IX contain a high affinity
 RT Ca2+-binding site that lacks gamma-carboxyglutamic acid.";
 RL J. Biol. Chem. 259:5698-5704(1984).
 RN [12]
 RP ERRATUM.

RA Morita T., Isaacs B.S., Esmon C.T., Johnson A.E.;
 RL J. Biol. Chem. 260:2583-2583(1985).
 RN [13]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE=86189947; PubMed=3009023;
 RA Bentley A.K., Rees D.J.G., Rizza C., Brownlee G.G.;
 RT "Defective propeptide processing of blood clotting factor IX caused
 RT by mutation of arginine to glutamine at position -4.";
 RL Cell 45:343-348(1986).
 RN [14]
 RP STRUCTURE OF CARBOHYDRATE ON SER-99.
 RX MEDLINE=90062160; PubMed=2511201;
 RA Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.,
 RA Shimonishi Y., Iwanaga S.;
 RT "Identification of a disaccharide (Xyl-Glc) and a trisaccharide
 RT (Xyl2-Glc) O-glycosidically linked to a serine residue in the first
 RT epidermal growth factor-like domain of human factors VII and IX and
 RT protein Z and bovine protein Z.";
 RL J. Biol. Chem. 264:20320-20325(1989).
 RN [15]
 RP STRUCTURE OF CARBOHYDRATE ON SER-99.
 RX MEDLINE=91344709; PubMed=2129367;
 RA Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;
 RT "A new trisaccharide sugar chain linked to a serine residue in the
 RT first EGF-like domain of clotting factors VII and IX and protein Z.";
 RL Adv. Exp. Med. Biol. 281:121-131(1990).
 RN [16]
 RP STRUCTURE OF CARBOHYDRATE ON SER-107.
 RX MEDLINE=92388094; PubMed=1517205;
 RA Nishimura H., Takao T., Hase S., Shimonishi Y., Iwanaga S.;
 RT "Human factor IX has a tetrasaccharide O-glycosidically linked to
 RT serine 61 through the fucose residue.";
 RL J. Biol. Chem. 267:17520-17525(1992).
 RN [17]
 RP PHOSPHORYLATION OF SER-114.
 RX MEDLINE=92388094; PubMed=1517205;
 RA Harris R.J., Papac D.I., Truong L., Smith K.J.;
 RT "Partial phosphorylation of serine-68 in EGF-1 of human factor IX.";
 RL (In) Abstracts of Xth international conference on methods in protein
 RL structure analysis, pp.50-50, Annecy (1996).
 RN [18]
 RP POST-TRANSLATIONAL MODIFICATIONS.
 RX MEDLINE=20575397; PubMed=11133752;
 RA Arruda V.R., Hagstrom J.N., Deitch J., Heiman-Patterson T.,
 RA Camire R.M., Chu K., Fields P.A., Herzog R.W., Couto L.B.,
 RA Larson P.J., High K.A.;
 RT "Posttranslational modifications of recombinant myotube-synthesized
 RT human factor IX.";
 RL Blood 97:130-138(2001).
 RN [19]
 RP STRUCTURE BY NMR OF 47-93.
 RX MEDLINE=95229607; PubMed=7713897;
 RA Freedman S.J., Furie B.C., Furie B., Baleja J.D.;
 RT "Structure of the metal-free gamma-carboxyglutamic acid-rich membrane
 RT binding region of factor IX by two-dimensional NMR spectroscopy.";
 RL J. Biol. Chem. 270:7980-7987(1995).
 RN [20]
 RP STRUCTURE BY NMR OF 47-93.
 RX MEDLINE=96032604; PubMed=7547952;
 RA Freedman S.J., Furie B.C., Furie B., Baleja J.D.;
 RT "Structure of the calcium ion-bound gamma-carboxyglutamic acid-rich
 RT domain of factor IX.";
 RL Biochemistry 34:12126-12137(1995).
 RN [21]
 RP STRUCTURE BY NMR OF 47-93.
 RX MEDLINE=96279169; PubMed=8663165;
 RA Freedman S.J., Blostein M.D., Baleja J.D., Jacobs M., Furie B.C.,
 RA Furie B.;
 RT "Identification of the phospholipid binding site in the vitamin K-
 RT dependent blood coagulation protein factor IX.";
 RL J. Biol. Chem. 271:16227-16236(1996).
 RN [22]
 RP STRUCTURE BY NMR OF 47-93.
 RX MEDLINE=97199336; PubMed=9047312;

RA Li L., Darden T.A., Freedman S.J., Furie B.C., Furie B., Baleja J.D.,
RA Smith H., Hiskey R.G., Pedersen L.G.;
RT "Refinement of the NMR solution structure of the
RT gamma-carboxyglutamic acid domain of coagulation factor IX using
RT molecular dynamics simulation with initial Ca²⁺ positions determined
RT by a genetic algorithm.";
RL Biochemistry 36:2132-2138(1997).
RN [23]
RP STRUCTURE BY NMR OF 91-133.
RX MEDLINE-91308127; PubMed-1854745;
RA Huang L.H., Cheng H., Pardi A., Tam J.P., Sweeney W.V.;
RT "Sequence-specific 1H NMR assignments, secondary structure, and
RT location of the calcium binding site in the first epidermal growth
RT factor like domain of blood coagulation factor IX.";
RL Biochemistry 30:7402-7409(1991).
RN [24]
RP STRUCTURE BY NMR OF 92-130.
RX MEDLINE-93284090; PubMed-1304885;
RA Baron M., Norman D.G., Harvey T.S., Handford P.A., Mayhew M.,
RA Tee A.G.D., Brownlee G.G., Campbell I.D.C.;
RT "The three-dimensional structure of the first EGF-like module of
RT human factor IX: comparison with EGF and IGF-alpha.";
RL Protein Sci. 1:81-90(1992).
RN [25]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 92-130.
RX MEDLINE-95330802; PubMed-7606779;
RA Rao Z., Handford P., Mayhew M., Knott V., Brownlee G.G., Stuart D.;
RT "The structure of a Ca(2+)-binding epidermal growth factor-like
RT domain: its role in protein-protein interactions.";
RL Cell 82:131-141(1995).
RN [26]
RP MOLECULAR PATHOLOGY OF HEMOPHILIA B.

Query Match 100.0%; Score 46; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
| | | | |
DB 376 LVDRATCLR 384

RESULT 9
FA9 RABIT STANDARD; PRT; 275 AA.
AC P16292;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Fragment).
GN F9.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90152675; PubMed-2303254;
RA Sarkar G., Koeberl D.D., Sommer S.S.;
RT "Direct sequencing of the activation peptide and the catalytic domain.
RT of the factor IX gene in six species.";
RL Genomics 6:133-143(1990).
CC -1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
CC PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
CC CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
CC IONS, PHOSPHOLIPIDS, AND FACTOR VIIIa.
CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC form factor Xa.
CC -1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIa, WHICH EXCISES THE
CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26234; AAA31251.1; -;
DR HSSP; P16293; LPFX.
DR MEROPS; S01.214; -;
DR InterPro; IPR001254; Ser protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00089; trypsin; 1;
DR SMART; SM00020; Tryp_SPC; 1;
DR PROSITE; PS00011; GLU CARBOXYLATION; PARTIAL.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
KW Hydrolase; Glycoprotein.
FT NON_TER 1
FT ACT_SITE 89 89 CHARGE RELAY SYSTEM.
FT ACT_SITE 137 137 CHARGE RELAY SYSTEM.
FT ACT_SITE 233 233 CHARGE RELAY SYSTEM.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 275 275
SQ SEQUENCE 275 AA; 30776 MW; FE364489AC76BE87 CRC64;

Query Match 91.3%; Score 42; DB 1; Length 275;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATCLR 9
| | | | |
DB 199 VDRATCLR 206

RESULT 10
VP40 ILTVT STANDARD; PRT; 586 AA.
AC P23984;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Capsid protein P40 [Contains: Capsid protein VP24 (Assemblin)
DE (protease) (EC 3.4.21.97); Capsid protein VP22A].
OS Infectious laryngotracheitis virus (strain Thorne V882) (ILTV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Bird herpesviruses;
OC Infectious laryngotracheitis-like viruses.
OX NCBI_TaxID=10344;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90301509; PubMed-2163526;
RA Griffin A.M.;
RT "The complete sequence of the capsid p40 gene from infectious
RT laryngotracheitis virus.";
RL Nucleic Acids Res. 18:3664-3664(1990).
RN [2]
RP SEQUENCE OF 1-516 FROM N.A.
RX MEDLINE-90218031; PubMed-2157797;
RA Griffin A.M., Boursnell M.E.;
RT "Analysis of the nucleotide sequence of DNA from the region of the
RT thymidine kinase gene of infectious laryngotracheitis virus;
RT potential evolutionary relationships between the herpesvirus
RT subfamilies.";
RL J. Gen. Virol. 71:841-850(1990).
CC -1- FUNCTION: VP22A IS A COMPONENT OF THE CAPSID CORE INVOLVED IN
CC PROCESSING AND PACKAGING OF PROGENY DNA. VP24 IS A PROTEASE

CC WHICH CAN PROTEOLYTICALLY CLEAVE ITSELF AND VP22A AT THE C-
 CC TERMINUS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Cleaves -Ala-|-Ser- and -Ala-|-Ala-bonds in
 CC the scaffold protein.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S21.
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP40.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00565; BAA00439.1; -.
 DR PIR; S13444; S13444.
 DR PIR; A43675; A43675.
 DR HSP; P16753; ICMV.
 DR MEROPS; S21.UPW; -.
 DR InterPro; IPR001847; Assemblin.
 DR Pfam; PF00716; Peptidase S21; 1.
 KW Coat protein; Hydrolase; Serine protease.
 FT CHAIN 1 239 COAT PROTEIN VP22A (PROTEASE).
 FT CHAIN 240 7558 COAT PROTEIN VP22A.
 FT PROPEP 7559 586 C-TERMINAL PEPTIDE.
 FT SITE 239 240 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
 FT SITE 558 559 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
 FT ACT_SITE 53 53 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 586 AA; 65335 MW; 0623A62933FFDB53 CRC64;
 Query Match 78.3%; Score 36; DB 1; Length 586;
 Best Local Similarity 66.7%; Pred. No. 7.5;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 Db 215 LMDRGTCCLK 223
 :|||:
 :|||:
 :|||:
 RESULT 11
 KROS HUMAN
 ID KROS HUMAN STANDARD; PRT; 2347 AA.
 AC P08922; Q15368;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase ROS precursor (BC 2.7.1.112)
 DE (c-ros-1).
 GN ROS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90280463; PubMed=2352949;
 RA Birchmeier C., O'Neill K., Riggs M., Wigler M.;
 RT "Characterization of ROS1 cDNA from a human glioblastoma cell line.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).
 RN [2]
 RP SEQUENCE OF 1790-2259 FROM N.A.
 RX MEDLINE=87064611; PubMed=3023956;
 RA Matsushime H., Wang L.-H., Shibuya M.;
 RT "Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma
 RT virus encodes for a transmembrane receptorlike molecule.";
 RL Mol. Cell. Biol. 6:3000-3004(1986).
 RN [3]
 RP SEQUENCE OF 1854-2245 FROM N.A.
 RX MEDLINE=87064625; PubMed=3785223;
 RA Birchmeier C., Birnbaum D., Waitches G., Fasano O., Wigler M.;

RT "Characterization of an activated human ros gene.";
 RL Mol. Cell. Biol. 6:3109-3116(1986).
 CC -!- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION
 CC FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34353; AAA60278.1; -.
 DR EMBL; M13599; AAA60277.1; -.
 DR EMBL; M13368; AAA60277.1; JOINED.
 DR EMBL; M13591; AAA60277.1; JOINED.
 DR EMBL; M13592; AAA60277.1; JOINED.
 DR EMBL; M13593; AAA60277.1; JOINED.
 DR EMBL; M13594; AAA60277.1; JOINED.
 DR EMBL; M13595; AAA60277.1; JOINED.
 DR EMBL; M13596; AAA60277.1; JOINED.
 DR EMBL; M13597; AAA60277.1; JOINED.
 DR EMBL; M13598; AAA60277.1; JOINED.
 DR EMBL; M13880; AAA6580.1; ALT TERM.
 DR PIR; A25223; TVHURS.
 DR PIR; A24421; TVHURS.
 DR HSP; P08631; IAD5.
 DR Genew; HGNC:10261; ROS1.
 DR MIM; 165020;
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR000033; Ldi_receptor_rec.
 DR InterPro; IPR002011; RTKinaseII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00041; fn3; 7.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00060; FN3; 5.
 DR SMART; SM00135; LY; 2.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; ATP-binding; Phosphorylation; Proto-oncogene;
 KW Signal.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 2347 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE
 FT ROS.
 FT DOMAIN 28 1859 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1860 1882 POTENTIAL.
 FT DOMAIN 1883 2347 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1945 2222 PROTEIN KINASE.
 FT NP_BIND 1951 1959 ATP (BY SIMILARITY).
 FT BINDING 1980 1980 ATP (BY SIMILARITY).
 FT MOD_RES 2114 2114 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 732 732 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 961 961 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1095 1095 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1211 1211 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1272 1272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1330 1330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1458 1458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1461 1461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1474 1474 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1565 1565 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1669 1669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1715 1715 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1738 1738 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1808 1808 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 2213 2213 N -> D (IN REF. 2 AND 3).
 FT CONFLICT 2228 2228 OC -> KS (IN REF. 2 AND 3).
 FT CONFLICT 2246 2246 EGDVCLNSDDIM -> KPDSEFSFRCTVN (IN REF. 2).
 SQ SEQUENCE 2347 AA; 263956 MW; E14F3DFD410C1D2A CRC64;
 Query Match 73.9%; Score 34; DB 1; Length 2347;
 Best Local Similarity 77.8%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRTATCLR 9
 |||:||||
 DB 309 LVDENHCLR 317

RESULT 12
 UL59 HCMVA STANDARD; PRT; 123 AA.
 AC P16746;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Hypothetical protein UL59.
 GN UL59.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Hornswell T., Hutchinson C.A. III, Kourzides T., Martignetti J.A.,
 RA Freddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169."
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X17403; CAA35374.1; -
 DR PIR; S09822; S09822.
 KW Hypothetical protein.
 SQ SEQUENCE 123 AA; 13945 MW; F730C581FA995562 CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 123;
 Best Local Similarity 66.7%; Pred. No. 6.8;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRTATCLR 9
 |||:||||
 DB 3 LVDQESCLR 11

RESULT 13
 CFAD_HUMAN STANDARD; PRT; 253 AA.
 AC P00746;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)
 DE (Properdin factor D) (Adipsin).
 GN DF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 8-253 FROM N.A.
 RX MEDLINE=92250520; PubMed=1374388;
 RA White R.T., Damm D., Hancock N., Rosen B.S., Lowell B.B., Usher P.,
 RA Flier J.S., Spiegelman B.M.;
 RT "Human adipsin is identical to complement factor D and is expressed
 RT at high levels in adipose tissue."
 RL J. Biol. Chem. 267:9210-9213(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Flier J.S., Spiegelman B.M., Rosen B.M.;
 RL Patent number WO9006365, 14-JUN-1990.
 RN [3]
 RP SEQUENCE OF 26-252.
 RX MEDLINE=85000441; PubMed=6383466;
 RA Niemann M.A., Shown A.S., Bennett J.C., Volanakis J.E.;
 RT "Amino acid sequence of human D of the alternative complement
 RT pathway."
 RL Biochemistry 23:2482-2486(1984).
 RN [4]
 RP PARTIAL SEQUENCE OF 26-252.
 RX MEDLINE=84108950; PubMed=6363133;
 RA Johnson D.M.A., Gagnon J., Reid K.B.M.;
 RT "Amino acid sequence of human factor D of the complement system.
 RT Similarity in sequence between factor D and proteases of non-plasma
 RT origin."
 RL FEBS Lett. 166:347-351(1984).
 RN [5]
 RP PARTIAL SEQUENCE OF 26-61 AND 194-220.
 RX MEDLINE=84256515; PubMed=6821372;
 RA Johnson D.M.A., Gagnon J., Reid K.B.M.;
 RT "Factor D of the alternative pathway of human complement.
 RT Purification, alignment and N-terminal amino acid sequences of the
 RT major cyanogen bromide fragments, and localization of the serine
 RT residue at the active site."
 RL Biochem. J. 187:863-874(1980).
 RN [6]
 RP PARTIAL SEQUENCE OF 26-82.
 RX MEDLINE=80145719; PubMed=6987665;
 RA Volanakis J.E., Shown A.S., Bennett J.C., Mole J.E.;
 RT "Partial amino acid sequence of human factor D: homology with serine
 RT proteases."
 RL Proc. Natl. Acad. Sci. U.S.A. 77:1116-1119(1980).
 RN [7]
 RP PARTIAL SEQUENCE OF 26-78.
 RX MEDLINE=81054886; PubMed=6776531;
 RA Davis A.E. III;
 RT "Active site amino acid sequence of human factor D."
 RL Proc. Natl. Acad. Sci. U.S.A. 77:4938-4942(1980).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=94118317; PubMed=8289289;

RA Narayana S.V.L., Carson M., El-Kabbani O., Kilpatrick J.M., Moore D.,
 RT Chen X., Bugg C.E., Volanakis J.E., Delucas L.J.;
 RT "Structure of human factor D. A complement system protein at 2.0-A
 resolution.";
 RL J. Mol. Biol. 235:695-708(1994).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=96025834; PubMed=7592653;
 RA Kim S., Narayana S.V., Volanakis J.E.;
 RT "Crystal structure of a complement factor D mutant expressing
 enhanced catalytic activity.";
 RL J. Biol. Chem. 270:24399-24405(1995).
 CC -!- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED
 CC WITH FACTOR C3B, ACTIVATING THE C3BBB COMPLEX, WHICH THEN BECOMES
 CC THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS
 CC HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY.
 CC -!- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-|-Lys) when in
 CC complex with C3b or with cobra venom factor (CVF).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- CAUTION: IN ADDITION TO THE CONFLICTS SHOWN IN THE FEATURE TABLE,
 CC REF.3 SEQUENCE HAD A PEPTIDE IN THE WRONG ORDER AND ANOTHER ONE
 CC MISSING.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M84526; AAA35527.1; ALT_INIT.
 DR PIR; A40197; DBHU.
 DR PDB; 1DFP; 25-FEB-98.
 DR PDB; 1DST; 11-JUL-96.
 DR PDB; 1DSU; 17-AUG-96.
 DR MEROPS; S01.191; -.
 DR Genew; HGNC:2771; DF.
 DR MIM; 134350; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Complement alternate pathway; Plasma; Hydrolase; Serine protease;
 KW Zymogen; Glycoprotein; Signal; 3D-structure.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 25 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 26 253 COMPLEMENT FACTOR D.
 FT ACT_SITE 66 66 CHARGE RELAY SYSTEM.
 FT ACT_SITE 112 112 CHARGE RELAY SYSTEM.
 FT ACT_SITE 208 208 CHARGE RELAY SYSTEM.
 FT DISULFID 51 67
 FT DISULFID 148 214
 FT DISULFID 179 195
 FT DISULFID 204 229
 FT CONFLICT 26 26
 FT CONFLICT 35 35
 FT CONFLICT 40 40
 FT CONFLICT 49 49
 FT CONFLICT 52 52
 FT CONFLICT 59 59
 FT CONFLICT 63 63
 FT CONFLICT 73 73
 FT CONFLICT 83 86
 FT CONFLICT 83 84
 FT CONFLICT 94 95
 FT CONFLICT 96 96
 FT CONFLICT 136 136
 FT CONFLICT 178 191
 I -> M (IN REF. 1).
 H -> F (IN REF. 6).
 M -> V (IN REF. 6).
 H -> E (IN REF. 4 AND 5).
 G -> A (IN REF. 1 AND 2).
 Q -> R (IN REF. 1 AND 2).
 S -> T (IN REF. 4).
 D -> G (IN REF. 4).
 HSLS -> THLP (IN REF. 3).
 HS -> ST (IN REF. 4).
 MISSING (IN REF. 4).
 D -> E (IN REF. 4).
 Q -> G (IN REF. 4).
 TCNRRTHDGAITE -> KCRLVDVL (IN REF. 4).

FT CONFLICT 243 243 S -> T (IN REF. 3).
 FT CONFLICT 250 250 S -> H (IN REF. 3).
 FT CONFLICT 250 250 MISSING (IN REF. 4).
 SQ SEQUENCE 253 AA; 27004 MW; BD553B70BD55C6AD CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 253;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 Db 173 VLDRATCNR 181
 :|||:|
 RESULT 14
 Y220 HUMAN
 ID Y220 HUMAN STANDARD; PRT; 546 AA.
 AC Q926T7; O43332; DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein KIAA0220.
 GN KIAA0220.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI TaxID=9606;
 RN [1]_SEQUENCE FROM N.A.
 RP TISSUE=Bone marrow;
 RX MEDLINE=97191544; PubMed=9039502;
 RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
 RA Ohata O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. VI.
 RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
 RT analysis of cDNA clones from cell line KG-1 and brain.";
 RL DNA Res. 3:321-329(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 RT human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D86974; BAAL3210.1; ALT_INIT.
 DR EMBL; AC003007; AAC31670.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 546 AA; 60996 MW; F09D0824566CEF71 CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 546;
 Best Local Similarity 66.7%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 Db 75 LVHRTGTCIR 83
 :|||:|
 RESULT 15
 CLCNR_HUMAN
 ID CLCNR_HUMAN STANDARD; PRT; 268 AA.

AC Q99895; Q9NUH5; O00765;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caldecrin precursor (EC 3.4.21.2) (Chymotrypsin C).
GN CTRC OR CLCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A., AND VARIANT TRP-80.
RP TISSUE=Pancreas;
RX MEDLINE=96221265; PubMed=8635596;
RA Tomomura A., Akiyama M., Itoh H., Yoshino I., Tomomura M., Nishii Y.,
RA Noikura T., Saeki T.;
RT "Molecular cloning and expression of human caldecrin.";
RL FEBS Lett. 386:26-28(1996).
RN [2]
RC SEQUENCE FROM N.A.
RP Coville G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE OF 17-268 FROM N.A.
RP TISSUE=Pancreas;
RA Sziegoleit A.;
RT "A human pancreatic chymotrypsin: biochemical and molecular
characterization.";
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RN CHARACTERIZATION.
RX MEDLINE=98207038; PubMed=9538241;
RA Yoshino-Yasuda I., Kobayashi K., Akiyama M., Itoh H., Tomomura A.,
RA Saeki T.;
RT "Caldecrin is a novel-type serine protease expressed in pancreas, but
its homologue, elastase IV, is an artifact during cloning derived
from caldecrin gene.";
RL J. Biochem. 123:546-554(1998).
CC -!- FUNCTION: HAS CHYMOTRYPSIN-TYPE PROTEASE ACTIVITY AND HYPOCALCEMIC
CC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Tyr|-Xaa,
CC Phe|-Xaa, Met|-Xaa, Trp|-Xaa, Gln|-Xaa, Asn|-Xaa.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S82198; AAB47104.2; ALT_SEQ.
DR EMBL; AL031283; CAB77355.1; -.
DR EMBL; Y13697; CAA74031.1; -.
DR HSSP; P00766; 1CHG.
DR MEROPS; S01.157; -.
DR Genew; HGNC:2523; CTRC.
DR MIM; 601405; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypain; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW Polymorphism.
FT SIGNAL. 1 16 POTENTIAL.
FT PROPEP 17 29 ACTIVATION PEPTIDE.
FT CHAIN 30 268 CALDECRIIN.

FT ACT_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 17 141 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 80 80 R -> W.
FT /FTID=VAR 010928.
FT CONFLICT 16 16 S -> T (IN REF. 1).
FT CONFLICT 52 52 N -> D (IN REF. 3).
SQ SEQUENCE 268 AA; 29484 MW; 460BF33B4A96516F CRC64;
Query Match 69.6%; Score 32; DB 1; Length 268;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
DB 180 VVDHATCSR 188

Search completed: May 6, 2003, 15:01:27
Job time : 10 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein , protein search, using sw model

Run on: May 6, 2003, 14:59:05 ; Search time 15.75 seconds
(without alignments)
54.934 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46

Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	271	2 I46580	factor IX - pig (f
2	46	100.0	274	2 I47078	coagulation factor
3	46	100.0	282	2 I84621	coagulation factor
4	46	100.0	285	2 I48144	coagulation factor
5	46	100.0	416	1 KFBO	coagulation factor
6	46	100.0	452	1 A30351	coagulation factor
7	46	100.0	459	2 JQ0419	coagulation factor
8	46	100.0	461	1 KFHU	coagulation factor
9	42	91.3	275	2 I46712	factor IX - rabbit
10	36	78.3	586	1 A43675	capsid protein p40
11	36	78.3	909	2 H87729	protein Y2JH5A.7
12	34	73.9	254	2 S49329	trypsin-like prote
13	34	73.9	434	2 AE3197	citrate synthase
14	34	73.9	2347	1 TVHURS	kinase-related pro
15	33	71.7	70	2 I40186	hypothetical prote
16	33	71.7	123	2 S09822	hypothetical prote
17	33	71.7	165	2 H83612	hypothetical prote
18	33	71.7	246	1 DBHU	complement factor
19	33	71.7	294	2 A49932	2,2',3-trihydroxyb
20	33	71.7	997	2 A60776	230k bullous pemph
21	32	69.6	162	2 H97253	deoxycyclylate de
22	32	69.6	268	2 S68825	pancreatic elastas
23	32	69.6	268	2 S68825	pancreatic elastas
24	32	69.6	323	2 C30338	DNA transport mach
25	32	69.6	343	2 E86310	protein Fli3.6 [im
26	32	69.6	502	2 AE3601	periplasmic oligop
27	32	69.6	538	2 E84863	hypothetical prote
28	32	69.6	547	1 S28904	intercellular adhe
29	32	69.6	688	2 S39491	protophyllid

30 32 69.6 863 1 S51789 VLDL receptor prec
31 32 69.6 2649 2 A40937 bullous pemphigoid
32 31 67.4 28 2 A44923 carboxypeptidase 3
33 31 67.4 124 2 S55042 11 kDa proteinase
34 31 67.4 130 2 A72720 hypothetical prote
35 31 67.4 162 2 A75443 hypothetical prote
36 31 67.4 286 2 F72599 hypothetical prote
37 31 67.4 308 2 F95903 conserved hypothet
38 31 67.4 378 2 AB3278 glutamate 5-kinase
39 31 67.4 417 2 E70660 hypothetical prote
40 31 67.4 421 2 B86973 probable membrane
41 31 67.4 431 2 A56822 synaptonemal compl
42 31 67.4 498 2 JQ2353 glycoprotein E pre
43 31 67.4 520 2 A13596 sugar transport Ar
44 31 67.4 543 2 B84398 hypothetical prote
45 31 67.4 555 2 T43847 DNA-directed RNA p

ALIGNMENTS

RESULT 1

I46580
factor IX - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 22-Jun-1999

C:Accession: I46580

R:Sarkar, G.; Koerberl, D.D.; Sommer, S.S.

Genomics 6, 133-143, 1990

A:Title: Direct sequencing of the activation peptide and the catalytic domain of the fa

A:Reference number: I46580; MUID:90152675; PMID:2303254

A:Accession: I46580

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-271 <SAR>

A:Cross-references: GB:M26235; NID:g164450; PIDN:AAA31031.1; PID:g164451

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology;

F:45-271/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 46; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.1; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 LVDRATCLR 9
DB 194 LVDRATCLR 202

RESULT 2

I47078

coagulation factor IXa (EC 3.4.21.22) - sheep (fragment)

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 22-Jun-1999

C:Accession: I47078

R:Sarkar, G.; Koerberl, D.D.; Sommer, S.S.

Genomics 6, 133-143, 1990

A:Title: Direct sequencing of the activation peptide and the catalytic domain of the fa

A:Reference number: I46580; MUID:90152675; PMID:2303254

A:Accession: I47078

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-274 <SAR>

A:Cross-references: GB:M26233; NID:g165878; PIDN:AAA31520.1; PID:g552419

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:49-274/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 46; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.1; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 LVDRATCLR 9
DB 194 LVDRATCLR 202

Db 198 LVDRATCLR 206

RESULT 3

184621

coagulation factor IXa (EC 3.4.21.22) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: I84621
R:Sarkar, G.; Koerberl, D.D.; Sommer, S.S.
Genomics 6, 133-143, 1990
A:Title: Direct sequencing of the activation peptide and the catalytic domain of the factor IXa
A:Reference number: I46580; MUID:90152675; PMID:2303254
A:Accession: I84621
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-282 <RES>
A:Cross-references: GB:M26247; NID:G204145; PIDN:AAA1162.1; PID:G204146
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:56-282/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 46; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

|||||

Db 205 LVDRATCLR 213

RESULT 4

148144

coagulation factor IXa (EC 3.4.21.22) - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I48144
R:Sarkar, G.; Koerberl, D.D.; Sommer, S.S.
Genomics 6, 133-143, 1990
A:Title: Direct sequencing of the activation peptide and the catalytic domain of the factor IXa
A:Reference number: I46580; MUID:90152675; PMID:2303254
A:Accession: I48144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-285 <RES>
A:Cross-references: GB:M26237; NID:G191260; PIDN:AAA37037.1; PID:G191261
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:59-285/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 46; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

|||||

Db 208 LVDRATCLR 216

RESULT 5

KF80

coagulation factor IXa (EC 3.4.21.22) precursor - bovine
N:Alternate names: Christmas factor
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Nov-1980 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
C:Accession: A14757; B20274; I45891; A00923
R:Katsayama, K.; Ericsson, L.H.; Enfield, D.L.; Waleh, K.A.; Neurath, H.; Davie, E.W.; Thromb. Haemostas. 66, 1-10, 1990
Proc. Natl. Acad. Sci. U.S.A. 76, 4990-4994, 1979
A:Title: Comparison of amino acid sequence of bovine coagulation factor IX (Christmas factor) with that of human factor IX
A:Reference number: A14757; MUID:80056619; PMID:231916
A:Accession: A14757
A:Molecule type: protein
A:Residues: 1-63, 'T', 65-416 <KAT>

R:McMullen, B.A.; Fujikawa, K.; Kistiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A:Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood coagulation factor IX
A:Reference number: A20274; MUID:83308813; PMID:6688526
A:Accession: B20274
A:Molecule type: protein
A:Residues: 59-63, 'X', 65-69 <MCM>
R:Choo, K.H.; Gould, K.G.; Rees, D.J.G.; Brownlee, G.G.
Nature 299, 178-180, 1982
A:Title: Molecular cloning of the gene for human anti-haemophilic factor IX.
A:Reference number: I45891; MUID:82272386; PMID:6287289
A:Accession: I45891
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 52-139 <CHO>
A:Cross-references: GB:J00007; NID:G163053; PIDN:AAA30520.1; PID:G163054
R:Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanaga, J. Biochem. 104, 867-868, 1988
A:Title: A new triaccharide sugar chain linked to a serine residue in bovine blood coagulation factor IX
A:Reference number: A44556; MUID:89213999; PMID:3149637
A:Contents: annotation
A>Note: structure and location of a carbohydrate covalently bound to Ser
C:Comment: Factor IX is activated by factor Xa, which excises the activation peptide p
C:Comment: The gamma-carboxyglutamic acid residues arise by posttranslational, vitamin K
C:Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with str
C:Function:
A:Description: catalyzes the proteolytic activation of coagulation factor X in the pres
A:Pathway: blood coagulation intrinsic pathway
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxygluta
F:1-146/Product: coagulation factor IXa light chain #status experimental <ALC>
F:1-45/Domain: Gla domain homology (fragment) <GLA>
F:51-82/Domain: EGF homology <EGL>
F:98-124/Domain: EGF homology <EG2>
F:147-181/Domain: activation peptide #status experimental <APT>
F:182-416/Product: coagulation factor IXa heavy chain #status experimental <AHC>
F:182-409/Domain: trypsin homology <TRY>
F:7, 8, 15, 17, 20, 21, 26, 27, 30, 33, 36, 40/Modified site: gamma-carboxyglutamic acid (Glu) #sta
F:18-23, 51-62, 56-71, 73-82, 88-99, 95-109, 111-124, 132-390, 207-223, 337-351, 362-390/Disulfide
F:53/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:64/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F:158, 168, 173, 261/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:222, 270, 366/Active site: His, Asp, Ser #status predicted
Query Match 100.0%; Score 46; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
|||||
Db 331 LVDRATCLR 339
RESULT 6
A30351
coagulation factor IXa (EC 3.4.21.22) precursor - dog
C:Species: Canis lupus familiaris (dog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A30351; I46201
R:Evans, J.P.; Watzke, H.H.; Ware, J.L.; Stafford, D.W.; High, K.A.
Blood 74, 207-212, 1989
A:Title: Molecular cloning of a cDNA encoding canine factor IX.
A:Reference number: A30351; MUID:89323338; PMID:2752110
A:Accession: A30351
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-452 <EVA>
A:Cross-references: GB:M21757; NID:9972719; PIDN:AAA75006.1; PID:G163948
R:Axelrod, J.H.; Read, M.S.; Brinkhouse, K.M.; Verma, I.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5173-5177, 1990
A:Title: Phenotypic correction of factor IX deficiency in skin fibroblasts of hemophilic
A:Reference number: I46201; MUID:90311364; PMID:2367529
A:Accession: I46201

A>Status: preliminary; translated from GB/EMBL/DBDJB

A:Molecule type: mRNA
 A:Residues: 1-452 <AXE>
 A:Cross-references: GB:M33826; NID:G163949; PIDN:AAA30844.1; PID:G163950
 C:Superfamily: coagulation factor IX; EGF homology; Gla domain homology; trypsin homology
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate
 F:1-21/Domain: signal sequence #status predicted <PRO>
 F:22-40/Domain: propeptide #status predicted <PRO>
 F:24-84/Domain: Gla domain homology <GLA>
 F:41-452/Product: coagulation factor IX #status predicted <MAT>
 F:90-121/Domain: EGF homology <EG1>
 F:127-163/Domain: EGF homology <EG2>
 F:218-445/Domain: trypsin homology <TRY>
 F:46-47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #
 F:57-62,90-101,95-110,112-121,127-138,134-148,150-163,171-326,243-259,373-387,398-426/Di
 F:258,306,402/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 46; DB 1; Length 452;

Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Gaps 0;

Matches 9; Conservative 0; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9

Db 367 LVDRATCLR 375

RESULT 7

JQ0419
 coagulation factor IXa (EC 3.4.21.22) precursor - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
 C:Accession: JQ0419; I49667
 R:Wu, S.M.; Stafford, D.W.; Ware, J.
 Gene 86, 275-278, 1990
 A>Title: Deduced amino acid sequence of mouse blood-coagulation factor IX.
 A:Reference number: JQ0419; MUID:90215309; PMID:2323576
 A:Accession: JQ0419
 A:Molecule type: mRNA
 A:Residues: 1-459 <MUS>
 A:Cross-references: GB:M23109; NID:G193317; PIDN:AAA37629.1; PID:G387158
 A:Experimental source: liver
 R:Sarkar, G.; Koerber, D.D.; Sommer, S.S.
 Genomics 6, 133-143, 1990
 A>Title: Direct sequencing of the activation peptide and the catalytic domain of the fac
 A:Reference number: I46580; MUID:90152675; PMID:2303254
 A:Accession: I49667

A>Status: preliminary; translated from GB/EMBL/DBDJB
 A:Molecule type: mRNA
 A:Residues: 168-362; Q, 364-387, I', 389-451 <RES>
 A:Cross-references: GB:M26236; NID:G193319; PIDN:AAA37630.1; PID:G193320
 C:Comment: This protein plays a critical role in blood coagulation.
 C:Superfamily: coagulation factor IX; EGF homology; Gla domain homology; trypsin homology
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate
 F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
 F:17-34/Domain: propeptide #status predicted <PRO>
 F:19-79/Domain: Gla domain homology <GLA>
 F:35-459/Product: coagulation factor IX #status predicted <MAT>
 F:85-116/Domain: EGF homology <EG1>
 F:122-158/Domain: EGF homology <EG2>
 F:225-452/Domain: trypsin homology <TRY>
 F:41,42,49,51,54,55,60,61,64,67,70,74/Modified site: gamma-carboxyglutamic acid (Glu) #
 F:52-57,85-96,90-105,107-116,122-133,129-143,145-158,166-333,250-266,380-394,405-433/Di
 F:265,313,409/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 46; DB 2; Length 459;

Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Gaps 0;

Matches 9; Conservative 0; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9

Db 374 LVDRATCLR 382

RESULT 8

KFHU
 coagulation factor IXa (EC 3.4.21.22) precursor [validated] - human
 N:Alternate names: antihemophilic factor B; Christmas factor
 C:Species: Homo sapiens (man)
 C:Date: 17-Dec-1982 #sequence_revision 30-Jun-1987 #text_change 15-Sep-2000
 C:Accession: A00922; A37570; A30511; A32989; A22673; A21337; A37546; A30623; A60486; A20
 R:Yoshitake, S.; Schach, B.G.; Foster, D.C.; Davie, E.W.; Kurachi, K.
 Biochemistry 24, 3736-3750, 1985
 A>Title: Nucleotide sequence of the gene for human factor IX (antihemophilic factor B).
 A:Reference number: A00922; MUID:86000558; PMID:2994716
 A:Accession: A00922
 A:Molecule type: DNA
 A:Residues: 1-461 <YOS>
 A:Cross-references: GB:K02402; NID:G182612; PIDN:AAB59620.1; PID:G182613
 R:Anson, D.S.; Choo, K.H.; Rees, D.J.G.; Giannelli, F.; Gould, K.; Huddleston, J.A.; Brc
 EMBO J. 3, 1053-1060, 1984
 A>Title: The gene structure of human anti-haemophilic factor IX.
 A:Reference number: A37570; MUID:84236100; PMID:6329734
 A:Accession: A37570
 A:Molecule type: DNA
 A:Residues: 1-461 <ANG>
 A:Cross-references: GB:K02048
 R:Reitsma, P.H.; Bertina, R.M.; Ploos van Amstel, J.K.; Riemens, A.; Briet, E.
 Blood 72, 1074-1076, 1988
 A>Title: The putative factor IX gene promoter in hemophilia B Leyden.
 A:Reference number: A30511; MUID:88327116; PMID:3416069
 A:Accession: A30511
 A:Molecule type: DNA
 A:Residues: 8-24 <REI>
 A:Cross-references: EMBL:X55008; NID:G311288; PIDN:CAB38245.2; PID:G4469253
 R:Koerber, D.D.; Bottama, C.D.K.; Buerstedde, J.M.; Sommer, S.S.
 Am. J. Hum. Genet. 45, 448-457, 1989
 A>Title: Functionally important regions of the factor IX gene have a low rate of polymor
 A:Reference number: A32989; MUID:89371752; PMID:2773937
 A:Accession: A32989
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 30-92 <KOB>
 R:McGraw, R.A.; Davis, L.M.; Noyes, C.M.; Lundblad, R.L.; Roberts, H.R.; Graham, J.B.; S
 Proc. Natl. Acad. Sci. U.S.A. 82, 2847-2851, 1985
 A>Title: Evidence for a prevalent dimorphism in the activation peptide of human coagulat
 A:Reference number: A22673; MUID:85190593; PMID:3857619
 A:Accession: A22673
 A:Molecule type: mRNA
 A:Residues: 1-193, 'T', 195-461 <MCG>
 A:Cross-references: GB:M1309; NID:G180552; PIDN:AAA52023.1; PID:G180553
 A:Note: The authors translated the codon ACA for residue 29 as Tyr
 R:Jaye, M.; de la Salle, H.; Schamber, F.; Balland, A.; Kohli, V.; Findeli, A.; Tolstoe
 Nucleic Acids Res. 11, 2325-2335, 1983
 A>Title: Isolation of a human anti-haemophilic factor IX cDNA clone using a unique 52-ba
 A:Reference number: A21337; MUID:83220788; PMID:6687940
 A:Accession: A21337
 A:Molecule type: mRNA
 A:Residues: 1-193, 'T', 195-461 <JAY>
 A:Cross-references: GB:J00137; NID:G182610; PIDN:AAA52763.1; PID:G182611
 R:Jagadeeswaran, P.; Lavelle, D.E.; Kaul, R.; Mohandas, T.; Warren, S.T.
 Somat. Cell Mol. Genet. 10, 465-473, 1984
 A>Title: Isolation and characterization of human factor IX cDNA: identification of Tag.I
 A:Reference number: A37546; MUID:84300526; PMID:6089357
 A:Accession: A37546
 A:Molecule type: mRNA
 A:Residues: 38-193, 'T', 195-326 <JAG>
 A:Cross-references: GB:M35672
 R:Kurachi, K.; Davie, E.W.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6461-6464, 1982
 A>Title: Isolation and characterization of a cDNA coding for human factor IX.
 A:Reference number: A30623; MUID:83065193; PMID:6959130
 A:Accession: A30623
 A:Molecule type: mRNA
 A:Residues: 1-12, 'S', 14-73, 'P', 75-82, 'K', 84-203, 'P', 205-216, 'G', 218-298, 'A', 299-356, 'A'
 A:Cross-references: GB:J00136; NID:G182608; PIDN:AAA98726.1; PMID:G182609
 A:Experimental source: liver

R;Tharakan, J.; Strickland, D.; Burgess, W.; Drohan, W.N.; Clark, D.B.
Vox Sang. 58, 21-29, 1990
A:Title: Development of an immunoaffinity process for factor IX purification.
A:Reference number: A60486; MUID:90194857; PMID:2316207
A:Accession: A60486
A:Molecule type: protein
A:Residues: 47-52, 'XX', 55-60, 'X', 62, 'XX', 65 <THA>
R;McMullen, B.A.; Fujikawa, K.; Kissel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A:Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood coagulation factor IX.
A:Reference number: A20274; MUID:83300813; PMID:6698526
A:Accession: A20274
A:Molecule type: protein
A:Residues: 105-109, 'X', 111-115 <MCM>
R;Balland, A.; Faure, T.; Carvalho, D.; Cordier, P.; Ulrich, P.; Fournet, B.; de la Salle, J.
Eur. J. Biochem. 172, 565-572, 1988
A:Title: Characterisation of two differently processed forms of human recombinant factor IX.
A:Reference number: S02527; MUID:88166735; PMID:3280312
A:Accession: S02527
A:Molecule type: protein
A:Residues: 29-63 <BAL>
A:Note: processed forms expressed in recombinant system
R;Jallat, S.; Perraud, F.; Dalemans, W.; Balland, A.; Diesterle, A.; Faure, T.; Meulien, E.
EMBO J. 9, 3295-3301, 1990
A:Title: Characterization of recombinant human Factor IX expressed in transgenic mice and its activation.
A:Reference number: S12058; MUID:9106024; PMID:2209546
A:Accession: S12058
A:Molecule type: mRNA; protein
A:Residues: 1-68 <JAL>
A:Note: processed forms expressed in recombinant system
R;Handford, P.A.; Baron, M.; Mayhew, M.; Willis, A.; Beesley, T.; Brownlee, G.G.; Campbell, J.
EMBO J. 9, 475-480, 1990
A:Title: The first EGF-like domain from human factor IX contains a high-affinity calcium binding site.
A:Reference number: S12377; MUID:90151623; PMID:2406129
A:Accession: S12377
A:Molecule type: protein
A:Residues: 92-130 <HAN>
A:Note: NMR detection of calcium binding by domain expressed in recombinant system
R;de la Salle, C.; Charmanlier, J.L.; Baas, M.J.; Schwartz, A.; Wiesel, M.L.; Grunbaum, A.; Thromb. Haemost. 70, 370-371, 1993
A:Title: A deletion located in the 3' non translated part of the factor IX gene responsible for a severe deficiency.
A:Reference number: S159612; MUID:94054330; PMID:8236150
A:Accession: S159612
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 444-461 <RES>
A:Cross-references: GB:S66752; NID:9439773; PIDN:AAB28588.1; PID:9439774
R;Stoflet, E.S.; Koeberl, D.D.; Sarkar, G.; Sommer, S.S.
Science 239, 491-494, 1988
A:Title: Genomic amplification with transcript sequencing.
A:Reference number: S159529; MUID:88127096; PMID:3340835
A:Accession: S159529
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 290-359 <RES>
A:Cross-references: GB:M19063; NID:G182622; PIDN:AAA52456.1; PID:G182623
R;Aarwala, K.L.; Kawabata, S.; Takao, T.; Murata, H.; Shimonishi, Y.; Nishimura, H.; Iwano, M.
Biochemistry 33, 5167-5171, 1994
A:Title: Activation peptide of human factor IX has oligosaccharides O-glycosidically linked to the active site.
A:Reference number: A54255; MUID:94227047; PMID:8172892
A:Accession: A54255
A:Molecule type: protein
A:Residues: 'D', 204, 'X', 206-211, 212, 'D', 214, 'X', 216-221, 'D' <AGA>
A:Note: the residues designated 'X' were determined to be threonine bound to carbohydrate
R;Di Scipio, R.G.; Kurachi, K.; Davie, E.W.
J. Clin. Invest. 61, 1528-1538, 1978
A:Title: Activation of human factor IX (Christmas factor).
A:Reference number: A18483; MUID:78194509; PMID:659613
R;McGraw, R.A.; Davis, L.M.; Noyes, C.W.; Graham, J.B.; Roberts, H.R.; Stafford, D.W.
Am. Soc. Hematol. Abstr. 64 (Suppl.1), 262a, 1984
A:Reference number: A37569
A:Accession: A37569
A:Contents: annotation

A:Note: 194-Thr was also found

J. Biol. Chem. 259, 5698-5704, 1984

A:Title: Derivatives of blood coagulation factor IX contain a high affinity Ca2+-binding site

A:Reference number: A37543; PMID:84185715; PMID:6425296

A:Contents: annotation; calcium binding

R:Morita, T.; Isaacs, B.S.; Esmon, C.T.; Johnson, A.E.

J. Biol. Chem. 260, 2583, 1985

A:Reference number: A37544

A:Contents: annotation; calcium binding, correction

R:Bentley, A.K.; Rees, D.J.G.; Rizza, C.; Brownlee, G.G.

Cell 45, 343-348, 1986

A:Title: Defective propeptide processing of blood clotting factor IX caused by mutation

A:Reference number: A37545; PMID:86189947; PMID:3009023

A:Contents: annotation; signal sequence cleavage site

R:Suehiro, K.; Kawabata, S.I.; Miyata, T.; Takeya, H.; Takamatsu, J.; Ogata, K.; Kamiya, J.

J. Biol. Chem. 264, 21257-21265, 1989

A:Title: Blood clotting factor IX B(M) Nagoya: substitution of arginine 180 by tryptophan

A:Reference number: A30622; PMID:90078229; PMID:2592373

A:Contents: annotation; sequence of mutant B(M) Nagoya

A:Note: carboxylation, glycosylation, and cleavage sites

R:Baron, M.; Norman, D.G.; Harvey, T.S.; Hanford, P.A.; Mayhew, M.; Tse, A.G.D.; Brown

submitted to the Brookhaven Protein Data Bank, November 1991

A:Reference number: A31252; PDB:1LXA

A:Contents: annotation; conformation by (1)H-NMR, residues 92-130

A:Note: recombinant form expressed in yeast

C:Comment: Factor IX is activated by factor Xla, which excises the activation peptide p

C:Comment: The gamma-carboxylglutamic acid residues arise by posttranslational, vitamin K

C:Comment: Calcium binds to the gamma-carboxylglutamic acid (Gla) residues and, with stro

C:Genetics:

A:Gene: GDB:F9

A:Cross-references: GDB:119900; OMIM:306900

A:Map position: XQ27.1-XQ27.2

A:Introns: 30/1; 84/2; 93/1; 131/1; 174/1; 241/3; 280/1

C:Function:

A:Description: catalyzes the proteolytic activation of coagulation factor X in the pres

A:Pathway: blood coagulation intrinsic pathway

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxylglut

F:1-28/Domain: signal sequence #status experimental <SIG>

F:29-46/Domain: propeptide #status experimental <PPT>

F:31-91/Domain: Gla domain homology <Gla>

F:97-128/Domain: EGF homology <EG1>

F:134-170/Domain: EGF homology <EG2>

F:192-226/Domain: activation peptide #status experimental <ACT>

F:227-461/Product: coagulation factor IXa heavy chain #status experimental <AHC>

F:227-454/Domain: trypsin homology <TRY>

F:53-54,61,63,66,67,72,73,76,79,82,86/Modified site: gamma-carboxylglutamic acid (Glu) #

F:64-69,97-108,102-117,119-128,134-145,141-155,157-170,178-335,252-268,382-396,407-435/D

F:99/Binding site: carboxylate (Ser) (covalent) #status experimental

F:110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

F:191-192/Cleavage site: Arg-Ala (coagulation factor Xla) #status experimental

F:203,213/Binding site: carboxylate (Asn) (covalent) #status experimental

F:205,215/Binding site: carboxylate (Thr) (covalent) #status experimental

F:226-227/Cleavage site: Arg-Val (coagulation factor Xla) #status experimental

Query Match 100.0%; Score 46; DB 1; Length 461;

Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9

Db 376 LVDRATCLR 384

RESULT 9

146712

factor IX - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 22-Jun-1999

C:Accession: 146712

R:Sarkar, G.; Koeberl, D.D.; Sommer, S.S.

Genomics 6, 133-143, 1990
A>Title: Direct sequencing of the activation peptide and the catalytic domain of the fac
A:Reference number: I46580; MUID:90152675; PMID:2303254
A:Accession: I46712
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-275 <SAR>
A:Cross-references: GB:M26234; NID:gl65020; PIDN:AAA31251.1; PID:gl65021
A:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
F:49-275/Domain: trypsin homology (fragment) <TRY>

Query Match 91.3%; Score 42; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 0.65; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Mismatches 0; Mismatches 0;

Qy 2 LVDRATCLR 9
|||||
Db 199 VDRATCLR 206
|||||

RESULT 10
A43675
capsid protein p40 - infectious laryngotracheitis virus
C:Species: infectious laryngotracheitis virus
C:Date: 30-Sep-1993 #sequence_revision 22-Oct-1999 #text_change 16-Jun-2000
A:Accession: S13444; A43675
R:Griffin, A.M.
Nucleic Acids Res. 18, 3664, 1990
A>Title: The complete sequence of the capsid p40 gene from infectious laryngotracheitis
A:Reference number: S13444; MUID:90301509; PMID:2163526
A:Accession: S13444
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-586 <NUC>
A:Cross-references: GB:D00565; NID:9221899; PIDN:BAA00439.1; PID:9221900
R:Griffin, A.M.; Boursnell, M.E.G.
J. Gen. Virol. 71, 841-850, 1990
A>Title: Analysis of the nucleotide sequence of DNA from the region of the thymidine kin
bramiles.
A:Reference number: A43675; MUID:90218031; PMID:2157797
A:Accession: A43675
A:Molecule type: DNA
A:Residues: 1-516, EF, <GRI>
A:Cross-references: GB:D00565
A:Experimental source: strain Thorne
C:Superfamily: varicella-zoster virus gene 33 protein
C:Keywords: capsid protein

Query Match 78.3%; Score 36; DB 1; Length 586;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9
|||||
Db 215 LVDRGTCCLK 223
|||||

RESULT 11
H87729
protein Y23H5A.7 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
A:Accession: H87729
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C. ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: H87729
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-909 <STO>

A:Cross-references: GB:chr_I; PIDN:AAC64633.1; PID:G3319446; GSPDB:GN00019; CESP:Y23H5A
C:Genetics:
A:Gene: Y23H5A.7
A:Map position: 1

Query Match 78.3%; Score 36; DB 2; Length 909;
Best Local Similarity 88.9%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9
|||||
Db 811 LVDRATCLR 819
|||||

RESULT 12
S49329
trypsin-like proteinase 1 (EC 3.4.21.-) precursor - imperfect fungus (Metarhizium anisop
C:Species: Metarhizium anisopliae
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 22-Jun-1999
A:Accession: JC4517; S49329
R:Smithson, S.L.; Paterson, I.C.; Bailey, A.M.; Screen, S.E.; Hunt, B.A.; Cobb, B.D.; C
Gene 166, 161-165, 1995
A>Title: Cloning and characterisation of a gene encoding a cuticle-degrading protease f
A:Reference number: JC4517; MUID:96105219; PMID:8529882
A:Accession: JC4517
A:Molecule type: DNA
A:Residues: 1-254 <SM2>
A:Cross-references: EMBL:X78875; NID:G556656; PIDN:CAA55477.1; PID:G556657
A:Experimental source: ME1
A>Note: Submitted to the EMBL Data Library, April 1994
C:Genetics:
A:Gene: try1
A:Introns: 86/3; 149/2
C:Superfamily: trypsin; trypsin homology
C:Keywords: endoplasmic reticulum; hydrolase; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:23-29/Domain: propeptide #status predicted <PRP>
F:30-254/Product: trypsin-like proteinase #status predicted <MAT>
F:30-250/Domain: trypsin homology <TRY>
F:70,112,209/Active site: His, Asp, Ser #status predicted

Query Match 73.9%; Score 34; DB 2; Length 254;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVDRATC 7
|||||
Db 169 VVDRATC 175
|||||

RESULT 13
AE3197
citrate synthase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid AT
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
A:Accession: AE3197
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE3197
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45995.1; PID:gi7743749; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: cisZ
A:Genome: plasmid

Query Match 73.9%; Score 34; DB 2; Length 434;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
 : : : : :
 DB 14 LIQRETCLR 22

RESULT 14
 TVHURS
 Kinase-related protein ros-1 precursor - human
 N/Alternate names: protein-tyrosine kinase mcf3 (activated ros-1)
 N/Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1
 C/Species: Homo sapiens (man)
 C/Date: 31-Dec-1988 #sequence_revision 07-Oct-1994 #text_change 11-Jun-1999
 C/Accession: A35512; A24421; A33081
 R/Birchmeier, C.; O'Neill, K.; Riggs, M.; Wigler, M.
 Proc. Natl. Acad. Sci. U.S.A. 87, 4799-4803, 1990
 A/Title: Characterization of ROS1 cDNA from a human glioblastoma cell line.
 A/Reference number: A35512; MUID:90280463; PMID:2352949
 A/Accession: A35512
 A/Molecule type: mRNA
 A/Residues: 1-2212, 'N', 2214-2227, 'QC', 2229-2347 <BIR>
 A/Cross-references: GB:M34353
 A/Experimental source: glioblastoma cell line SW-1088
 R/Matsushima, H.; Wang, L.H.; Shibuya, M.
 Mol. Cell. Biol. 6, 3000-3004, 1986
 A/Title: Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma virus encode
 A/Reference number: A25223; MUID:87064611; PMID:3023956
 A/Accession: A25223
 A/Molecule type: DNA
 A/Residues: 1790-2245, 'KPDSSPSSFRCTVN' <MA2>
 A/Cross-references: GB:M13368
 A/Experimental source: placenta
 A/Note: the differences after residue 2245 result from the authors' misinterpretation of
 R/Birchmeier, C.; Birnbaum, D.; Waitches, G.; Fasano, O.; Wigler, M.
 Mol. Cell. Biol. 6, 3109-3116, 1986
 A/Title: Characterization of an activated human ros gene.
 A/Reference number: A24421; MUID:87064625; PMID:3785223
 A/Accession: A24421
 A/Molecule type: mRNA
 A/Residues: 1854-2261, 'A', 2263-2347 <BI2>
 A/Cross-references: GB:M13880; NID:9337482; PIDN:AAA36580.1; PID:9337483
 A/Experimental source: tumor cells
 A/Note: the mcf3 oncogene was formed by DNA rearrangement involving fusion of at least b
 C/Genetics:
 A/Gene: GDB:ROS1
 A/Cross-references: GDB:120351; OMIM:165020
 A/Map position: 6q22-6q22
 A/Introns: 1853/1; 1861/1; 1926/2; 1980/3; 2002/2; 2045/3; 2078/2; 2145/2; 2190/2
 C/Suprafamily: Kinase-related protein ros; LDL receptor YWTD-containing repeat homology;
 C/Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 F1-36/Domain: signal sequence #status predicted <SIG>
 F137-2347/Product: kinase-related protein ROS1 #status predicted <MAT>
 F137-1859/Domain: extracellular #status predicted <EXT>
 F1335-378/Domain: LDL receptor YWTD-containing repeat homology <YWI>
 F1466-503/Domain: LDL receptor YWTD-containing repeat homology <YWA>
 F1715-757/Domain: LDL receptor YWTD-containing repeat homology <YW2>
 F1758-798/Domain: LDL receptor YWTD-containing repeat homology <YW3>
 F1799-838/Domain: LDL receptor YWTD-containing repeat homology <YW4>
 F1843-888/Domain: LDL receptor YWTD-containing repeat homology <YW5>
 F1893-933/Domain: LDL receptor YWTD-containing repeat homology <YW6>
 F1532-1574/Domain: LDL receptor YWTD-containing repeat homology <YW7>
 F1860-1883/Domain: transmembrane #status predicted <TMN>
 F1884-2347/Domain: intracellular #status predicted <INT>
 F11943-2222/Domain: protein kinase homology <KIN>
 F12951-1959/Region: protein kinase ATP-binding motif
 F152,114,123,324,352,471,607,628,706,714,732,935,961,1015,1087,1090,1211,1272,1330,1458,
 F1980/Active site: Lys #status predicted
 F1210,2114,2115/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 73.9%; Score 34; DB 1; Length 2347;
 Best Local Similarity 77.8%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
 : : : : :
 DB 309 LVDEAHCLR 317

RESULT 15
 I40186
 hypothetical protein I169b - Bacteroides fragilis
 C/Species: Bacteroides fragilis
 C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999
 C/Accession: I40186; S41295
 R/Firrh, S.; Heggoud, A.; Reyeser, G.; Sebal, M.
 Microbiology 141, 927-935, 1995
 A/Title: Plasmids pIP419 and pIP421 from Bacteroides: 5-nitroimidazole resistance genes
 A/Reference number: I40184; MUID:95291442; PMID:7773395
 A/Accession: I40186
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-70 <RES>
 A/Cross-references: EMBL:X76949; NID:9440383; PIDN:CAAS4271.1; PID:9440385

Query Match 71.7%; Score 33; DB 2; Length 70;
 Best Local Similarity 62.5%; Pred. No. 12;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 8
 : : : : :
 DB 57 MLDKATCL 64

Search completed: May 6, 2003, 15:03:27
 Job time : 17.75 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 15:01:10 ; Search time 18 Seconds
(without alignments)
43.143 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46

Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 8628685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : * Published Applications AA:*

1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	181	10	US-09-864-761-46600
2	46	100.0	461	9	US-10-132-829-5
3	46	100.0	461	10	US-09-884-901-3
4	41	89.1	415	10	US-09-118-748-2
5	33	71.7	101	9	US-10-012-542-351
6	33	71.7	228	10	US-09-821-255-2
7	32	69.6	44	12	US-10-001-843-188
8	32	69.6	83	10	US-09-864-761-35244
9	32	69.6	252	9	US-10-117-323-38
10	32	69.6	434	10	US-09-753-436-120
11	32	69.6	547	9	US-10-047-542-65
12	32	69.6	547	10	US-09-753-436-1
13	32	69.6	2125	10	US-09-919-172-29
14	31	67.4	24	9	US-10-012-542-362
15	31	67.4	223	10	US-09-870-162A-13
16	31	67.4	338	10	US-09-886-055-205
17	31	67.4	382	10	US-09-731-872-312
18	31	67.4	429	9	US-10-012-542-366
19	31	67.4	466	10	US-09-731-872-311

20	31	67.4	470	9	US-10-012-542-365
21	31	67.4	873	9	US-10-167-264-2
22	30	65.2	10	9	US-09-572-404B-1740
23	30	65.2	215	10	US-09-925-297-660
24	30	65.2	363	9	US-10-045-815-2
25	30	65.2	374	9	US-10-255-817-10
26	30	65.2	736	9	US-10-045-815-4
27	29	63.0	80	10	US-09-764-847-626
28	29	63.0	80	10	US-09-864-761-34304
29	29	63.0	193	10	US-09-864-761-34304
30	29	63.0	282	9	US-10-167-555-2
31	29	63.0	300	9	US-10-224-414-2
32	29	63.0	326	9	US-10-108-915-46
33	29	63.0	459	9	US-09-759-595-3
34	29	63.0	460	9	US-10-001-254-16
35	29	63.0	460	9	US-10-001-254-28
36	29	63.0	460	9	US-09-759-595-1
37	29	63.0	909	9	US-10-078-770-190
38	29	63.0	909	12	US-10-078-729-168
39	28	60.9	10	9	US-09-572-404B-1749
40	28	60.9	36	9	US-10-083-357-1143
41	28	60.9	57	10	US-09-864-761-36792
42	28	60.9	70	12	US-10-107-371-1
43	28	60.9	75	9	US-10-104-755-12
44	28	60.9	77	9	US-09-792-793A-14
45	28	60.9	119	9	US-10-141-965-2

ALIGNMENTS

RESULT 1

US-09-864-761-46600
; Sequence 46600, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenhang
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46600
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033403.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EST HUMAN HIT: T28608.1, EVALUATE 3.00e-36
; OTHER INFORMATION: SWISSPROT HIT: P00740, EVALUATE 1.00e-107
US-09-864-761-46600

Query Match      100.0%; Score 46; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
Db 96 LVDRATCLR 104

RESULT 2
US-10-132-829-5
; Sequence 5, Application US/10132829
; Publication No. US2003004982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-PAL170
; CURRENT APPLICATION NUMBER: US/10/132,829
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-5

Query Match      100.0%; Score 46; DB 9; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
Db 376 LVDRATCLR 384

RESULT 3
US-09-884-901-3
; Sequence 3, Application US/09884901
; Patent No. US20020076798A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Carol
; APPLICANT: Kay, Mark
; TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods of Use
; FILE REFERENCE: U0FW-1-17396
; CURRENT APPLICATION NUMBER: US/09/884,901
; CURRENT FILING DATE: 2001-06-18
```

```
; PRIOR APPLICATION NUMBER: US 60/212,902
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-901-3

Query Match      100.0%; Score 46; DB 10; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
Db 376 LVDRATCLR 384

RESULT 4
US-09-118-748-2
; Sequence 2, Application US/09118748A
; Patent No. US20020031799A1
; GENERAL INFORMATION:
; APPLICANT: Stafford, Darrel W.
; APPLICANT: Chang, Jinli
; TITLE OF INVENTION: Factor IX Antihemophilic Factor with Increased Clotting
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 5470-183
; CURRENT APPLICATION NUMBER: US/09/118,748A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,571
; EARLIER FILING DATE: 1997-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-118-748-2

Query Match      89.1%; Score 41; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 8
Db 330 LVDRATCLR 337

RESULT 5
US-10-012-542-351
; Sequence 351, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029PI
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-351

Query Match 71.7%; Score 33; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DRATCL 8
Db 33 DRATCL 38

RESULT 6

US-09-821-255-2
; Sequence 2, Application US/09821255
; Patent No. US20020081293A1
; GENERAL INFORMATION:
; APPLICANT: Michael S.C. Fung
; APPLICANT: Bill N.C. Sun
; APPLICANT: Cecily R.Y. Sun
; TITLE OF INVENTION: Inhibitors of Complement Activation
; FILE REFERENCE: 98-2A
; CURRENT APPLICATION NUMBER: US/09/821,255
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/075,328
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: 09/253,689
; PRIOR FILING DATE: 1999-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: human
US-09-821-255-2

Query Match 71.7%; Score 33; DB 10; Length 228;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
Db 148 VLDATCLR 156

RESULT 7

US-10-001-843-188
; Sequence 188, Application US/10001843
; Patent No. US20020132255A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
; FILE REFERENCE: DEX-0267
; CURRENT APPLICATION NUMBER: US/10/001,843
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,992
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 188
; LENGTH: 44

; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-843-188

Query Match 69.6%; Score 32; DB 12; Length 44;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCL 8
Db 7 LVERAVCL 14

RESULT 8

US-09-864-761-35244
; Sequence 35244, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35244
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033378.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: P37743, EVALUATE 6.90e-01
; OTHER INFORMATION: EST_HUMAN HIT: AUI41284.1, EVALUATE 5.00e-31
US-09-864-761-35244

```

```

Query Match 69.6%; Score 32; DB 10; Length 83;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 VDRATCLR 9
DB 10 VGRSTCLR 17

```

RESULT 9

```

US-10-117-323-38
; Sequence 38, Application US/10117323
; Publication No. US20030054993A1
; GENERAL INFORMATION:
; APPLICANT: Rancourt, Derrick E.
; APPLICANT: O'Sullivan, Colleen M.
; TITLE OF INVENTION: Implantation Serine Proteinases
; FILE REFERENCE: 033337-005
; CURRENT APPLICATION NUMBER: US/10/117,323
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/281,724
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/294,736
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/350,962
; PRIOR FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Mouse
US-10-117-323-38

```

```

Query Match 69.8%; Score 32; DB 9; Length 252;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 LVDRATCLR 9
DB 164 VVDHATCSR 172

```

RESULT 10

```

US-09-753-436-120
; Sequence 120, Application US/09753436
; Patent No. US20010029293A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemary
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,436
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/382,289
; FILING DATE:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Joseph A., Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 33282
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-753-436-120

```

```

Query Match 69.6%; Score 32; DB 10; Length 434;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VDRATC 7
DB 292 IDRATC 297

```

RESULT 11

```

US-10-047-542-65
; Sequence 65, Application US/10047542
; Patent No. US20020168367A1
; GENERAL INFORMATION:
; APPLICANT: LARRICK, JAMES W.
; APPLICANT: WYCOFF, KEITH L.
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
; FILE REFERENCE: 030905.0004.CIP1
; CURRENT APPLICATION NUMBER: US/10/047,542
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US01/13932
; PRIOR FILING DATE: 2001-04-28
; PRIOR APPLICATION NUMBER: 60/200,298
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 101

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-047-542-65

Query Match 69.6%; Score 32; DB 9; Length 547;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
:|||||
Db 402 IDRATC 407

RESULT 12
US-09-753-436-1
; Sequence 1, Application US/09753436
; Patent No. US20010029293A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,436
; FILING DATE:

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/382,289
; FILING DATE:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Joseph A., Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 33282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 547 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 30..547
US-09-753-436-1

Query Match 69.6%; Score 32; DB 10; Length 547;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
:|||||
Db 402 IDRATC 407

RESULT 13
US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 3774181CDI
US-09-919-172-29

Query Match 69.6%; Score 32; DB 10; Length 2125;
Best Local Similarity 77.8%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
:|||||
Db 1946 LVDRITALR 1954

RESULT 14
US-10-012-542-362
; Sequence 362, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 362
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-362

Query Match 67.4%; Score 31; DB 9; Length 24;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRTCL 8
Db 12 LQDRASCL 19

RESULT 15
US-09-870-162A-13
; Sequence 13, Application US/09870162A
; Patent No. US20020042118A1
; GENERAL INFORMATION:
; APPLICANT: Breinig, Sabine
; APPLICANT: Fuchs, Georg
; TITLE OF INVENTION: Phenol Induced Proteins of Thauera aromatica
; FILE REFERENCE: BC1006 US DIV
; CURRENT APPLICATION NUMBER: US/09/870,162A
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 09/516914
; PRIOR FILING DATE: 2000-03-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Thauera aromatica
US-09-870-162A-13

Query Match 67.4%; Score 31; DB 10; Length 223;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRTCLR 9
Db 46 LITRAGCLR 54

Search completed: May 6, 2003, 15:04:52
Job time : 19.3333 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:30 ; Search time 14.4 Seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46
Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2.6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2.6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2.6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2.6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2.6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2.6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	40	2	US-08-797-842-11
2	46	100.0	40	2	US-08-797-842-12
3	46	100.0	235	4	US-08-944-483-48
4	46	100.0	415	1	US-08-073-531B-1
5	46	100.0	415	1	US-08-295-411-2
6	46	100.0	415	2	US-08-955-471-2
7	46	100.0	415	2	US-08-766-288-1
8	46	100.0	415	5	PCT-US92-10242-2
9	46	100.0	461	3	US-08-742-877-2
10	46	100.0	461	4	US-09-053-871A-21
11	46	100.0	461	6	5521070-2
12	33	71.7	228	4	US-08-944-483-44
13	33	71.7	238	6	5223425-5
14	33	71.7	250	6	5223425-4
15	33	71.7	253	6	5223425-8
16	32	69.6	239	4	US-08-944-483-61
17	32	69.6	268	1	US-08-270-584A-2
18	32	69.6	268	1	US-08-568-031-2
19	32	69.6	268	2	US-08-765-192-2
20	32	69.6	268	2	US-08-966-319-2
21	32	69.6	268	3	US-09-153-304-2
22	32	69.6	268	3	US-09-199-793-2
23	32	69.6	434	2	US-08-487-113D-120
24	32	69.6	434	2	US-08-720-420A-120
25	32	69.6	518	4	US-09-240-915-8
26	32	69.6	518	4	US-09-591-435-8
27	32	69.6	547	1	US-08-314-615-1

28	32	69.6	547	1	US-08-314-362-1	Sequence 1, Appli
29	32	69.6	547	1	US-08-473-981A-6	Sequence 6, Appli
30	32	69.6	547	1	US-08-433-010-1	Sequence 1, Appli
31	32	69.6	547	1	US-08-482-882-1	Sequence 1, Appli
32	32	69.6	547	2	US-08-483-389-1	Sequence 1, Appli
33	32	69.6	547	2	US-08-487-113D-1	Sequence 1, Appli
34	32	69.6	547	2	US-08-473-503-1	Sequence 1, Appli
35	32	69.6	547	2	US-08-483-932-1	Sequence 1, Appli
36	32	69.6	547	2	US-08-474-087-6	Sequence 6, Appli
37	32	69.6	547	2	US-08-720-420A-1	Sequence 1, Appli
38	32	69.6	547	3	US-08-714-017-1	Sequence 1, Appli
39	32	69.6	547	3	US-08-863-790-1	Sequence 1, Appli
40	32	69.6	547	3	US-08-475-680-1	Sequence 1, Appli
41	32	69.6	547	4	US-08-296-749-1	Sequence 1, Appli
42	31	67.4	77	4	US-09-134-001C-3500	Sequence 3500, Ap
43	31	67.4	222	6	5223425-6	Patent No. 5223425
44	31	67.4	223	4	US-09-516-914-13	Sequence 13, Appli
45	31	67.4	846	1	US-08-149-103-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-797-842-11
; Sequence 11, Application US/08797842
; Patent No. 5932706
; GENERAL INFORMATION:

APPLICANT: Mertens, Koenraad et al

TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
TITLE OF INVENTION: their use for isolating intact protein, haemostatic composi
TITLE OF INVENTION: of proteolytic cleavage products of the protein

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Michaelson and Wallace

STREET: Parkway 109 Office Center, 328 Newman Springs

CITY: Red Bank

STATE: New Jersey

COUNTRY: USA

ZIP: 07701

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette

COMPUTER: IBM PC

OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11

SOFTWARE: Microsoft Word for Windows 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/797,842

FILING DATE: 10-Feb-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/381,891

FILING DATE: February 8, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Michaelson, Peter L.

REGISTRATION NUMBER: 30090

REFERENCE/DOCKET NUMBER: Stitching-5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)530-6671

TELEFAX: (908)530-6584

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 40 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-797-842-11

Query Match 100.0%; Score 46; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
| | | | |
DB 23 LVDRATCLR 31

RESULT 2

US-08-797-842-12
; Sequence 12, Application US/08797842
; Patent No. 5932706
; GENERAL INFORMATION:
; APPLICANT: Mertens, Koenraad et al
; TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
; TITLE OF INVENTION: their use for isolating intact protein, haemostatic composition
; TITLE OF INVENTION: of proteolytic cleavage products of the protein
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michaelson and Wallace
; STREET: Parkway 109 Office Center, 328 Newman Springs
; STREET: Road, P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11
; SOFTWARE: Microsoft Word for Windows 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,842
; FILING DATE: 10-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/381,891
; FILING DATE: February 8, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Stitching-5
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-797-842-12

Query Match 100.0%; Score 46; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
| | | | |
DB 12 LVDRATCLR 20

RESULT 3

US-08-944-483-48
; Sequence 48, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.

; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6193.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-48

Query Match 100.0%; Score 46; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
| | | | |
DB 150 LVDRATCLR 158

RESULT 4

US-08-073-531B-1
; Sequence 1, Application US/08073531B
; Patent No. 5621039
; GENERAL INFORMATION:
; APPLICANT: Hallahan, et al.
; TITLE OF INVENTION: Factor IX - Polymeric Conjugates
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GALCANO & BURKE
; STREET: 300 Rabro Drive
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS 6.0
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/073,531B

;; FILING DATE: June 8, 1993
;; CLASSIFICATION: 525
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: -
;; FILING DATE: -
;; ATTORNEY/AGENT INFORMATION:
;; NAME: GALGANO & BURKE
;; REGISTRATION NUMBER: 30,735
;; REFERENCE/DOCKET NUMBER: 128-7
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 582-6161
;; TELEFAX: (516) 582-6191
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 415 Amino Acids
;; TYPE: Amino Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Unknown to applicant
;; MOLECULE TYPE: -
;; HYPOTHETICAL: -
;; ANTI-SENSE: -
;; ORIGINAL SOURCE:
;; ORGANISM:
;; INDIVIDUAL ISOLATE: -
;; CELL TYPE: -
;; IMMEDIATE SOURCE:
;; LIBRARY:
;; CLONE:
;; PUBLICATION INFORMATION:
;; AUTHORS:
;; TITLE:
;; JOURNAL:
;; VOLUME:
;; ISSUE:
;; PAGES:
;; DATE:
;; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 415.

US-08-073-531B-1

Query Match 100.0%; Score 46; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9
Db 330 LVDRATCLR 338

RESULT 5
US-08-295-411-2
; Sequence 2, Application US/08295411
; Patent No. 5679639
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Metters, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,471,
; FILING DATE:

;; APPLICATION NUMBER: US/08/295,411
;; FILING DATE: 22-AUG-1994
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/793,989
;; FILING DATE: 18-NOV-1991
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fitting, Thomas
;; REGISTRATION NUMBER: 34,163
;; REFERENCE/DOCKET NUMBER: TSRI263.0C1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619-554-2937
;; TELEFAX: 619-554-6312
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 415 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 1..145
;; OTHER INFORMATION: /note= "Factor IX Light Chain"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 146..180
;; OTHER INFORMATION: /note= "Factor IX Activation"
;; OTHER INFORMATION: Peptide"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 181..415
;; OTHER INFORMATION: /note= "Factor IX Heavy Chain"
US-08-295-411-2.

Query Match 100.0%; Score 46; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVDRATCLR 9
Db 330 LVDRATCLR 338

RESULT 6
US-08-955-471-2
; Sequence 2, Application US/08955471
; Patent No. 5968751
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Metters, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,471,
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/295,411
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSI263.0C1
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note= "Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note= "Factor IX Activation"
OTHER INFORMATION: Peptide"
FEATURE:
NAME/KEY: Region
LOCATION: 181..415
OTHER INFORMATION: /note= "Factor IX Heavy Chain"
US-08-953-471-2

Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 330 LVDRATCLR 338

RESULT 7
US-08-766-288-1
Sequence 1, Application US/08766288
Patent No. 5969040
GENERAL INFORMATION:
APPLICANT: Hallahan, et al.
TITLE OF INVENTION: Factor IX - Polymeric Conjugates
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: GALGANO & BURKE
STREET: 300 Rabro Drive
CITY: Hauppauge
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS 6.0
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,288
FILING DATE:
CLASSIFICATION: 525
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,531
FILING DATE: June 8, 1993
ATTORNEY/AGENT INFORMATION:
NAME: GALGANO & BURKE

REGISTRATION NUMBER: 30,735
REFERENCE/DOCKET NUMBER: 128-7 (DIV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 582-6161
TELEFAX: (516) 582-6191
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 Amino Acids
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown to applicant
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE:
CELL TYPE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
PUBLICATION INFORMATION:
AUTHORS:
TITLE: JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 415.
US-08-766-288-1
Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
DB 330 LVDRATCLR 338
RESULT 8
PCT-US92-10242-2
Sequence 2, Application PC/TUS9210242
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
APPLICANT: Mesters, Rolf
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
TITLE OF INVENTION: for Inhibiting Coagulation
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESSEE: Research Institute
STREET: 10666 North Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10242
FILING DATE: 19921118
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,989
FILING DATE: 18-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas

Search completed: May 6, 2003, 15:04:05
Job time : 16.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:44 ; Search time 40.95 Seconds
(without alignments)
29.286 Million cell updates/sec

Title: US-09-851-422b-1

Perfect score: 46

Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	181	22	Peptide #3514 enco
2	46	100.0	182	22	Human factor IX (h
3	46	100.0	412	16	Human Factor-IX.
4	46	100.0	415	14	Factor IX (IX). H
5	46	100.0	454	16	Human Factor-IX.
6	46	100.0	456	5	Part of the sequen
7	46	100.0	456	5	Sequence encoded b
8	46	100.0	461	6	Sequence of human
9	46	100.0	461	6	Sequence of human
10	46	100.0	461	6	Sequence of human

11	46	100.0	461	11	AA05393	Mutant human facto
12	46	100.0	461	19	AAW40284	Human Factor IX pr
13	46	100.0	461	21	AAW97295	Human clotting fac
14	46	100.0	461	22	AAE10828	Human wild-type fa
15	46	100.0	461	22	AAE60281	Human factor IX (h
16	46	100.0	461	23	AAO21524	Protein of human f
17	46	100.0	461	23	AA050568	Human Factor IX.
18	46	100.0	462	12	AA010868	Recombinant human
19	41	89.1	415	20	AA03203	Amino acid sequenc
20	37	80.4	367	22	AAU47611	Propionibacterium
21	34	73.9	63	20	AA011751	Human 5' EST secre
22	34	73.9	63	21	AA000629	Human secreted pro
23	34	73.9	158	12	AA011721	ENV93/HTLV-1-IIIB'
24	33	71.7	96	22	AAU14883	Novel bone marrow
25	33	71.7	101	21	AA086427	Human gene 31-enco
26	33	71.7	138	20	AAU31395	HEV peptide z12-or
27	33	71.7	205	22	AAU47245	Propionibacterium
28	33	71.7	212	20	AA028590	Human Factor D. H
29	33	71.7	250	11	AA05421	Human adipsin/D en
30	33	71.7	250	11	AA05772	Human adipsin gene
31	33	71.7	345	22	AB063738	Drosophila melanog
32	33	71.7	575	22	AA039439	Human polypeptide
33	33	71.7	581	22	AB011366	Human KIAA0220 pro
34	33	71.7	581	22	AA041225	Human polypeptide
35	33	71.7	877	22	AB062624	Drosophila melanog
36	33	71.7	997	12	AA015345	Mouse bullous pemph
37	33	71.7	1374	22	AB012335	Novel human diagno
38	32	69.6	12	14	AA030472	ICAM-3 peptide fra
39	32	69.6	38	22	AA090306	Human immune/haema
40	32	69.6	83	22	AB023163	Peptide #2014 enco
41	32	69.6	83	22	AB034537	Peptide #2043 enco
42	32	69.6	83	22	AB019946	Protein #1945 enco
43	32	69.6	83	22	AA055323	Human brain expres
44	32	69.6	83	22	AA067720	Human bone marrow
45	32	69.6	83	22	AA015530	Peptide #1964 enco

ALIGNMENTS

RESULT 1

ABB36008
ID ABB36008 standard; Peptide; 181 AA.
AC ABB36008;
XX
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3514 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DB;
XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 28643; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 181 AA;
Query Match 100.0%; Score 46; DB 22; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
| | | | |
Db 96 LVDRATCLR 104
RESULT 2
AAB60289
ID AAB60289 standard; Protein; 182 AA.
XX
AC AAB60289;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human factor IX (hFIX) exon-encoded fragment, SEQ ID NO:12.
XX
XX Age-related gene regulation; liver-specific; gene expression;
KW human factor IX; hFIX; AFS'; age-regulatable expression construct;
KW antiseize therapy; gene therapy; thrombosis; cardiovascular disease;
KW diabetes; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
KW osteoarthritis; dementia.
XX
OS Homo sapiens.
XX
XX WO200075279-A2.
XX
PD 14-DEC-2000.
XX
PP 06-JUN-2000; 2000WO-US15728.
XX
PP 09-JUN-1999; 99US-0328925.
XX
PR (UNMI) UNIV MICHIGAN.
XX
PA Kurachi K, Kurachi S;
XX
PI WPI; 2001-061708/07.
XX
DR N-PSDB; AAF54018.
XX
XX New regulatory elements that control age-related gene expression,
PT useful in gene therapy and for reducing Factor IX expression -
XX
XX Disclosure; Fig 8E; 225pp; English.
XX
CC The invention relates to nucleic acid sequences which regulate gene
CC expression in an age-related manner and/or in a liver-specific manner.
CC The invention identifies regions of the human factor IX (hFIX) gene, and
CC a region of the human protein C (hPC) gene, which are age-related
CC regulatory sequences. The hFIX age-related regulatory sequences are
CC designated AFS' (AAF54016) and AFS' (AAF54017) and are found in the 5'
CC UTR (at position 2164-2165 of AAF54016) and 3' UTR (at position
CC 34383-35655 of AAF54018) respectively. These elements act synergistically

CC to increase hFIX levels over the lifespan of an individual; however, they
CC can independently exert effects on hFIX mRNA in an age-related manner.
CC with AFS' acting to stabilise hFIX mRNA, and AFS' acting to increase hFIX
CC mRNA levels, over time. AFS' also directs liver-specific expression. The
CC hPC gene age-related regulatory sequence is found in the 5' UTR
CC (AAF54081), and contains two PEA-3 (polyoma virus activator 3) elements
CC 5'-GAGGAAA-3' and 5'-CAGGAG-3'. The age-related regulatory sequences of
CC the invention, along with their homologues, variants and fragments, may
CC be used in the construction of recombinant expression vectors for the
CC expression of a desired sequence in an age-related fashion in a host
CC cell. Preferred target genes for expression in such age-regulatable
CC expression vectors include those encoding proteins involved in blood
CC coagulation (e.g., the pro-coagulants factor IX and factor VIII, and the
CC anti-coagulants protein C and antithrombin III), human
CC alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
CC luciferase. Preferred promoters for use in such age-regulatable
CC expression vectors include the human factor IX promoter, the T7 promoter,
CC the T3 promoter and the SP6 promoter. The expression vectors of the
CC invention may be used in gene therapy to provide age-related and/or
CC liver-specific expression of target genes. Age-regulatable constructs may
CC be used in the treatment of such age-related conditions such as
CC thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
CC Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
CC Specifically, they may be used to express factor IX antisense mRNA in the
CC treatment of thrombotic conditions associated with the natural
CC age-related rise in factor IX expression. Transgenic cells or animals
CC that contain vectors of the invention are useful as models of these
CC diseases, in screening for potential therapeutic agents and for studying
CC normal processes such as ageing and gene expression. Fragments and
CC homologues of age-related regulatory sequences, are useful as probes to
CC detect, isolate or identify other such sequences in samples. The present
CC sequence represents a fragment of hFIX.
XX
SQ Sequence 182 AA;
Query Match 100.0%; Score 46; DB 22; Length 182;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
| | | | |
Db 97 LVDRATCLR 105
RESULT 3
AAR64266
ID AAR64266 standard; Protein; 412 AA.
XX
AC AAR64266;
XX
DT 18-JUL-1995 (first entry)
XX
DE Human Factor-IX.
XX
XX Factor-IX; polymer conjugate; hemophilia-B; blood-clotting;
KW coagulant; polyethylene glycol; PEG.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 1..46
FT /label= Gla domain
FT /note= "gamma carboxy-glutamate-rich domain"
FT Modified-site 6
FT /label= OTHER
FT Modified-site 7
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site 14
FT /label= OTHER
FT Modified-site 16
FT /note= "gamma-carboxyglutamic acid"


```

XX PS Disclosure; Page 126-128; 149pp; English.
XX CC The PC polypeptides indicated in the Features Table inhibit
XX CC coagulation (they prevent binding of serine protease to natural
XX CC substrates), esp. when admin. to give an intravascular blood
XX CC concn. of 0.1-100 (pref. 0.5-10) microM.
XX CC NB: Sequences corresp. to SEQ ID NO 6, 7, 8 and 9 are described
XX CC in the specification but have not yet been added to the SEQUENCE
XX CC LISTING.
XX SQ Sequence 415 AA;
    Query Match 100.0%; Score 46; DB 14; Length 415;
    Best Local Similarity 100.0%; Pred. No. 0.96; 0; Indels 0; Gaps 0;
    Matches 9; Conservative 0; Mismatches 0;
QY 1 LVDRATCLR 9
Db 330 LVDRATCLR 338
    |||||
    |||||

RESULT 5
AAR67710
ID AAR67710 standard; Protein; 454 AA.
XX AC AAR67710;
XX XX
XX DT 20-JUL-1995 (first entry)
XX DE Human Factor-IX.
XX KW Factor-IX; blood-clotting; hemophilia B; gene therapy;
XX KW adenovirus; vector.
XX OS Homo sapiens.
XX PN WO9429471-A.
XX PD 22-DEC-1994.
XX PF 13-APR-1994; 94WO-US04075.
XX PR 10-JUN-1993; 93US-0074920.
XX PR 25-MAR-1994; 94US-0218335.
XX XX (GENE-) GENETIC THERAPY INC.
XX PI Connelly S, Kaleko M, Smith T;
XX DR WPI; 1995-036495/05.
XX DR N-PSDB; AAQ76017.
XX XX
XX PT New adenoviral vectors for treatment of haemophilia - contg. a
XX PT DNA sequence encoding a clotting factor, partic. Factor VIII or
XX PT Factor IX
XX PS Disclosure; Fig. 11A-11B; 116pp; English.
XX XX
XX CC Human Factor-IX was used to construct recombinant adenovirus vectors
XX CC that produced therapeutic levels of the clotting factor when
XX CC administered to an animal host, potentially providing hemophilia B
XX CC gene therapy.
XX SQ Sequence 454 AA;
    Query Match 100.0%; Score 46; DB 16; Length 454;
    Best Local Similarity 100.0%; Pred. No. 1.1;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
Db 369 LVDRATCLR 377
    |||||
    |||||

```

```

RESULT 6
AAP40178
ID AAP40178 standard; Protein; 456 AA.
XX AC AAP40178;
XX DT 11-FEB-1992 (first entry)
XX DE Part of the sequence of human factor IX.
XX KW Christmas disease; therapy; haemophilia; factor IX; blood clotting;
XX KW diagnosis.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..41
XX FT Protein 42..456
XX PN WO8400560-A.
XX PD 16-FEB-1984.
XX PF 03-AUG-1983; 83WO-GB00191.
XX PR 06-MAY-1983; 83GB-0012491.
XX PR 04-AUG-1982; 82GB-0022485.
XX XX
XX PA (NATR ) NATIONAL RES DEV CORP.
XX PA (BROW/) BROWNLEE G G.
XX PI Brownlee G, Choo KH;
XX DR WPI; 1984-049331/08.
XX DR N-PSDB; AAN40142.
XX XX
XX PT Recombinant DNA cloning vehicles - useful in prodn. of factor IX
XX PT polypeptide and of diagnostic probes for Christmas disease
XX PS Example; Fig 9; 71pp; English.
XX CC The inventors claim DNA molecules comprising part or all of the
XX CC human factor IX DNA. The invention also includes cDNA derived from
XX CC human factor IX RNA. Specifically claimed are: recombinant DNA (the
XX CC phage present in clone lambda HIX-1) deposited as NCIB No. 11749;
XX CC Recombinant DNA in which the cloning vehicle is the modified pAT 153
XX CC plasmid present in E.coli NCIB No. 11747; Recombinant DNA in which
XX CC the bovine factor IX DNA sequence is contained in the recombinant
XX CC DNA transformed into E.coli to form a clone deposited as NCIB No.
XX CC 11748.
XX SQ Sequence 456 AA;
    Query Match 100.0%; Score 46; DB 5; Length 456;
    Best Local Similarity 100.0%; Pred. No. 1.1;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
Db 371 LVDRATCLR 379
    |||||
    |||||

RESULT 7
AAP40222
ID AAP40222 standard; Protein; 456 AA.
XX AC AAP40222;
XX DT 13-FEB-1992 (first entry)
XX DE Sequence encoded by part of the sequence of human factor IX cDNA.

```

XX Haemophilia; Christmas disease; diagnosis; treatment.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Peptide 1..41
 FT Protein /label= signal
 FT 42..456
 XX GB2125409-A.
 XX PD 07-MAR-1984.
 XX PF 03-AUG-1983; 83GB-0020975.
 XX PR 16-MAY-1983; 83GB-0012490.
 XX PR 04-AUG-1982; 82GB-0022486.
 XX PR 03-AUG-1983; 83GB-0020975.
 XX (NATR) NATIONAL RES DEV CORP.
 XX Brownlee GG, Choo KH;
 XX WPI; 1984-057898/10.
 XX N-PSDB; AAN40177.
 XX Prodn. of artificial human factor IX - by use of recombinant DNA
 XX sequences for host transformation and cultivation
 XX Example; Fig 9; 49pp; English.
 XX The inventors claim a recombinant DNA having a human factor IX
 XX sequence pref. at least 50 nucleotides long, esp. 75-27000
 XX nucleotides. A cloning vector contg. foreign DNA is also claimed.
 XX The foreign sequence pref. includes the whole of an exon sequence of
 XX the human factor IX genome. The cloning vehicle may be a modified
 XX pAT 153 plasmid. Also claimed is a labelled diagnostic probe
 XX comprising a DNA molecule having a single- or double-stranded probe
 XX sequence of 15 to 10000 nucleotides long Factor IX DNA sequence.
 XX SQ Sequence 456 AA;
 Query Match 100.0%; Score 46; DB 5; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 DB 371 LVDRATCLR 379
 RESULT 8
 AAP50311
 ID AAP50311 standard; Protein; 461 AA.
 XX AC AAP50311;
 XX DT 03-SEP-1991 (first entry)
 XX DE Sequence of human factor IX, encoded by DNA FIX.
 XX KW Blood clotting; haemophilia B.
 XX OS Homo sapiens.
 XX W08505125-A.
 XX PD 21-NOV-1985.
 XX PF 09-MAY-1984; 84WO-0007125.
 XX PR 09-MAY-1984; 84FR-0007125.

XX (TRAN-) TRANSGENE SA.
 PA (JAYE/) JAYE M.
 XX PI Jaye M, De La Salle H, Tolstoshev P, Lecocq JP;
 XX WPI; 1985-302926/48.
 DR N-PSDB; AAN50362.
 XX Cloning and expression vector for factor nine - for transforming
 PT bacteria, yeast or mammalian cells
 XX Disclosure; Fig 3; 98pp; French.
 XX The inventors claim a vector which comprises a DNA SQ coding for
 CC factor IX (FIX) and elements providing expression of this SQ in host
 CC cells. The prepn. of DNA FIX is as follows. Polyadenylated mRNA is
 CC isolated from human liver, used to make ss-DNA which is digested
 CC with S1 and the fragments over 1 Kb cloned in pBR 322. Clones are
 CC selected by hybridisation with a 52-mer probe corresp. to the bovine
 CC factor IX SQ (AAN50362).
 XX SQ Sequence 461 AA;
 Query Match 100.0%; Score 46; DB 6; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 DB 376 LVDRATCLR 384
 RESULT 9
 AAP50302
 ID AAP50302 standard; Protein; 461 AA.
 XX AC AAP50302;
 XX DT 03-SEP-1991 (first entry)
 XX DE Sequence of human factor IX.
 XX KW Blood clotting; haemophilia B; vaccinia vector; cowpox vector.
 XX OS Homo sapiens.
 XX W08505376-A.
 XX PD 05-DEC-1985.
 XX PF 29-MAY-1985; 85WO-1104408.
 XX PR 05-OCT-1984; 84FR-0015294.
 XX PR 22-MAY-1984; 84FR-0007959.
 XX (TRAN-) TRANSGENE SA.
 PA (DSAL/) DE LA SALLE H.
 XX PI De La Salle H, Drillien R, Altenburger W, Tolstoshev P, Lecocq JP;
 XX WPI; 1985-316884/50.
 DR N-PSDB; AAN50351.
 XX Vector for expressing factor 9 in vertebrate cells - contg.
 PT poxvirus genome in which factor 9 gene is inserted
 XX Example; Fig 2; 59pp; French.
 XX The factor IX coding SQ (AAN50351) is first inserted into a non-
 CC essential region of viral DNA, cloned in a bacterial plasmid. A
 CC double-reciprocal recombination procedure is then used to transfer
 CC the insert into the viral genome where it is propagated and

XX 18-DEC-1997.
 XX 11-JUN-1997; 97WO-EP03027.
 XX 06-JUL-1996; 96EP-0110959.
 XX 11-JUN-1996; 96EP-0109288.
 XX 22-JUN-1996; 96EP-0110109.
 XX (BOEF) BOEHRINGER MANNHEIM GMBH.
 XX Hopfner K, Kopetzki E;
 XX WPI; 1998-052304/05.
 XX DR N-PSDB; AAV10463.
 XX Non-glycosylated, truncated forms of factor IX family protein with
 PT serine protease activity - used to screen for specific modulators
 PT and to assay factor IXa
 XX Disclosure; Fig 4; 49pp; German.
 XX This sequence represents a human factor IX protease. This protein is used
 CC in the construction of a novel non-glycosylated protein and truncated
 CC and zymogen forms of this protein, which have serine protease activity.
 CC The protein is composed of various domains from a factor IX family
 CC protein, namely a catalytic domain (CD) N-terminally bound to a
 CC zymogen-activating domain (ZAD), N-terminally bound to an EGF1 and/or
 CC EGF2 domain (EGF = epidermal growth factor-like domain). Such proteins
 CC are used to identify activators/inhibitors of factor IX family proteins
 CC (potentially useful as regulators of coagulation, fibrinolysis and
 CC homeostasis). The protein in zymogen form is also useful in assays for
 CC detecting factor IXa activity in aqueous solution (specifically in body
 CC fluids). The protein can be used to produce co-crystals with protease
 CC variants or inhibitors for X-ray structural analysis and drug modelling
 CC and as restriction proteases in biotechnology. These truncated proteins
 CC have the same specificity as factor IX family proteases and can be
 CC produced in prokaryotes in a form that allows production of active enzyme
 CC by conversion to native form and enzymatic cleavage.
 XX SQ Sequence 461 AA;
 Query Match 100.0%; Score 46; DB 19; Length 461;
 Best Local Similarity 100.0%; Pred. NO. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 DB 376 LVDRATCLR 384
 RESULT 13
 ID AAY97295 standard; Protein; 461 AA.
 AC AAY97295;
 XX 03-JAN-2001 (first entry)
 XX Human clotting factor IX.
 XX Recombinant DNA; gene therapy; hormone responsive element;
 KW transgene; HRE; haemophilia; clotting factor IX; vaccine;
 KW regulation; breast cancer; ovarian cancer; prostate cancer;
 KW von Willebrand disease; cystic fibrosis; hormone; receptor; human;
 KW blood.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200049147-A1.
 PN 24-AUG-2000.
 XX

PF 18-FEB-2000; 200WO-EP01368.
 XX 19-FEB-1999; 99DE-1007099.
 PR 19-FEB-1999; 99US-0120848.
 XX (THER-) THERAGENE BIOMEDICAL LAB GMBH.
 PA Hauser-funke C;
 XX WPI; 2000-549273/50.
 XX DR N-PSDB; AAA53835.
 XX Novel nucleic acid construct useful in gene therapy comprising an
 PT hormone responsive element and transgene in which the hormone
 PT responsive element is not functionally linked to the transgene
 XX Disclosure; Page 81-83; 100pp; English.
 XX New nucleic acid constructs are described which comprise an hormone
 CC responsive element (HRE) and a transgene (T). Alternatively the
 CC nucleic acid construct, comprises at least one HRE and a transgene,
 CC where one of the HREs is not functionally linked to the transgene;
 CC The constructs can be used to up-regulate or down-regulate target
 CC genes and for the delivery of vaccines. The constructs preferably
 CC comprise a transgene which encodes a protein which is lacking in a
 CC variety of genetic disorders or involved in conditions related in
 CC inappropriate responses to hormones, for example hormone-dependent
 CC cancers such as breast, ovarian, and endometrial cancers and prostate
 CC cancer. The transgene may also be used to replace a defective gene
 CC resulting in such genetic disorders as haemophilia, von Willebrand
 CC disease, and cystic fibrosis. Vectors comprising these constructs
 CC where the transgene is human clotting factor IX can be used for
 CC treating blood clotting disorders such as haemophilia A or B on
 CC administration to an organism or to a cellular system. The constructs
 CC have applications in gene therapy for treating haemophilia when the
 CC transgene encodes a clotting factor such as clotting factor IX. The
 CC advantage of this system is that the hormone-hormone receptor complex
 CC contains a hormone receptor that becomes activated after binding of
 CC its specific hormone. The hormone receptor in the activated state is
 CC able to recognise and bind to its specific hormone responsive
 CC element. The presence of the hormone responsive element on the
 CC nucleic acid carrying a transgene encourages binding of a
 CC hormone-hormone receptor complex. Thus the activated hormone receptor
 CC acts as a link between the nucleic acid carrying the transgene and the
 CC hormone known to interact with the cell membrane.
 XX SQ Sequence 461 AA;
 Query Match 100.0%; Score 46; DB 21; Length 461;
 Best Local Similarity 100.0%; Pred. NO. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 DB 376 LVDRATCLR 384
 RESULT 14
 ID AAE10828 standard; Protein; 461 AA.
 XX AAE10828;
 XX 18-DEC-2001 (first entry)
 XX Human wild-type factor IX protein.
 XX Human; haemostatic; coagulant; blood clotting factor; factor VIII;
 KW factor IX; therapy; haemophilia A.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200170968-A2.
 PN

XX 27-SRP-2001.
 PD
 XX
 PF 21-MAR-2001; 2001WO-EP03220.
 XX
 PR 22-MAR-2000; 2000BP-0106225.
 PR 08-MAY-2000; 2000US-203249P.
 XX
 PA (OCTA-) OCTAGENE GMBH.
 XX
 PI Hauser C, Hoerster A, Schroeder C, Lehnerer M;
 XX
 XX WPI; 2001-590175/66.
 DR N-PSDB; AAD18173.
 DR
 XX Recombinantly producing human blood coagulation factors VIII and IX for
 PT use in treating hemophilia -
 XX
 XX Claim 21; Page 73-74; 104pp; English.
 XX
 CC The present invention relates to an improved method for the production
 CC of recombinant human blood clotting factors, especially factors VIII and
 CC IX, using an immortalised cell line stably expressing viral transcription
 CC activators and carrying a vector comprising a promoter and a sequence
 CC encoding the blood coagulation factor. The factor VIII protein or a gene
 CC transfer vector is used in the preparation of agents for treating
 CC haemophilia, especially haemophilia A. The present sequence is human
 CC wild-type factor IX protein encoded by vector pIGF36hyg.
 XX
 SQ Sequence 461 AA;
 Query Match 100.0%; Score 46; DB 22; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 |||||
 DB 376 LVDRATCLR 384
 RESULT 15
 AAB60281
 ID AAB60281 standard; Protein; 461 AA.
 XX
 AC AAB60281;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human factor IX (hFIX) protein.
 XX
 KW Age-related gene regulation; liver-specific; gene expression;
 KW human factor IX; hFIX; ABE; ABE3; age-regulatable expression construct;
 KW antisense therapy; gene therapy; thrombosis; cardiovascular disease;
 KW diabetes; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
 KW osteoarthritis; dementia.
 XX
 OS Homo sapiens.
 XX
 PN WO200075279-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 06-JUN-2000; 2000WO-US15728.
 XX
 PR 09-JUN-1999; 99US-0328925.
 XX
 XX (UNMI) UNIV MICHIGAN.
 XX
 PI Kurachi K, Kurachi S;
 XX
 XX WPI; 2001-061708/07.
 DR N-PSDB; AAF54018.
 XX

PT New regulatory elements that control age-related gene expression,
 PT useful in gene therapy and for reducing Factor IX expression -
 XX
 PS Disclosure; Fig 8A-E; 225pp; English.
 XX
 CC The invention relates to nucleic acid sequences which regulate gene
 CC expression in an age-related manner and/or in a liver-specific manner.
 CC The invention identifies regions of the human factor IX (hFIX) gene, and
 CC a region of the human protein C (hPC) gene, which are age-related
 CC regulatory sequences. The hFIX age-related regulatory sequences are
 CC designated ABE3' (AAF54016) and ABE3' (AAF54017) and are found in the 5'
 CC UTR (at position 2164-2165 of AAF54018) and 3' UTR (at position
 CC 34383-35655 of AAF54018) respectively. These elements act synergistically
 CC to increase hFIX levels over the lifespan of an individual; however, they
 CC can independently exert effects on hFIX mRNA in an age-related manner.
 CC with ABE3' acting to stabilise hFIX mRNA, and ABE3' acting to increase hFIX
 CC mRNA levels, over time. ABE3' also directs liver-specific expression. The
 CC hPC gene age-related regulatory sequence is found in the 5' UTR
 CC (AAF54081), and contains two PEA-3 (polyoma virus activator 3) elements
 CC (5'-GAGGAAA-3' and 5'-CAGGAG-3'). The age-related regulatory sequences of
 CC the invention, along with their homologues, variants and fragments, may
 CC be used in the construction of recombinant expression vectors for the
 CC expression of a desired sequence in an age-related fashion in a host
 CC cell. Preferred target genes for expression in such age-regulatable
 CC expression vectors include those encoding proteins involved in blood
 CC coagulation (e.g., the pro-coagulants factor IX and factor VIII, and the
 CC anti-coagulants protein C and antithrombin III), human
 CC alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
 CC luciferase. Preferred promoters for use in such age-regulatable
 CC expression vectors include the human factor IX promoter, the T7 promoter,
 CC the T3 promoter and the SP6 promoter. The expression vectors of the
 CC invention may be used in gene therapy to provide age-related and/or
 CC liver-specific expression of target genes. Age-regulatable constructs may
 CC be used in the treatment of such age-related conditions such as
 CC thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
 CC Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
 CC Specifically, they may be used to express factor IX antisense mRNA in the
 CC treatment of thrombotic conditions associated with the natural
 CC age-related rise in factor IX expression. Transgenic cells or animals
 CC that contain vectors of the invention are useful as models of these
 CC diseases, in screening for potential therapeutic agents and for studying
 CC normal processes such as ageing and gene expression. Fragments and
 CC homologues of age-related regulatory sequences, are useful as probes to
 CC detect, isolate or identify other such sequences in samples. The present
 CC sequence represents hFIX.
 XX
 SQ Sequence 461 AA;
 Query Match 100.0%; Score 46; DB 22; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LVDRATCLR 9
 |||||
 DB 376 LVDRATCLR 384

Search completed: May 6, 2003, 15:01:01
 Job time : 42.95 secs

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:45 ; Search time 14.2 Seconds
(without alignments)
58.041 Million cell updates/sec

Title: US-09-851-422B-2

Perfect score: 20
Sequence: 1 DRAT 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	25	10 P82952	P82952 prunus dulc
2	20	100.0	27	12 Q91KR6	Q91KR6 hepatitis c
3	20	100.0	31	1 Q8X254	Q8X254 halobacteri
4	20	100.0	45	3 Q9URG8	Q9URG8 rhizomucor
5	20	100.0	46	2 Q9S604	Q9S604 escherichia
6	20	100.0	48	16 Q8XVT7	Q8XVT7 ralestonia s
7	20	100.0	49	10 Q9LJS3	Q9LJS3 arabidopsis
8	20	100.0	53	13 Q90XF3	Q90XF3 coturnix co
9	20	100.0	60	5 Q9U4L5	Q9U4L5 tetrahymena
10	20	100.0	63	12 Q65240	Q65240 african swi
11	20	100.0	63	17 Q9HQ75	Q9HQ75 halobacteri
12	20	100.0	64	9 Q9AYU9	Q9AYU9 lactococcus
13	20	100.0	64	10 Q8WQJ3	Q8WQJ3 oryza sativ
14	20	100.0	64	16 Q9CFP1	Q9CFP1 lactococcus
15	20	100.0	66	4 Q9NTN6	Q9NTN6 homo sapien
16	20	100.0	67	16 Q9L6P7	Q9L6P7 salmonella

17	20	100.0	67	16 Q8X3Y5	Q8X3Y5 escherichia
18	20	100.0	73	16 Q9L2J1	Q9L2J1 streptomyce
19	20	100.0	80	4 Q9NU58	Q9NU58 homo sapien
20	20	100.0	83	4 Q9NR86	Q9NR86 homo sapien
21	20	100.0	84	16 Q98P82	Q98P82 rhizobium 1
22	20	100.0	85	9 Q980B4	Q980B4 mycobacteri
23	20	100.0	86	17 Q9HL23	Q9HL23 thermoplasm
24	20	100.0	87	16 Q9PD28	Q9PD28 xylella fas
25	20	100.0	87	16 Q9XW88	Q9XW88 ralestonia s
26	20	100.0	91	2 Q06656	Q06656 streptomyce
27	20	100.0	91	12 Q39862	Q39862 human astro
28	20	100.0	91	16 Q9KDH8	Q9KDH8 bacillus ha
29	20	100.0	92	2 Q88153	Q88153 rhizobium e
30	20	100.0	94	4 Q96JUR8	Q96JUR8 homo sapien
31	20	100.0	97	10 Q8VZ93	Q8VZ93 arabidopsis
32	20	100.0	99	10 Q8SB16	Q8SB16 oryza sativ
33	20	100.0	100	11 Q54845	Q54845 mus musculu
34	20	100.0	102	4 Q8TB31	Q8TB31 homo sapien
35	20	100.0	102	16 Q9KY18	Q9KY18 streptomyce
36	20	100.0	103	10 Q8S076	Q8S076 oryza sativ
37	20	100.0	103	12 Q96819	Q96819 human astro
38	20	100.0	104	10 Q9AUP1	Q9AUP1 oryza sativ
39	20	100.0	108	5 Q9V4G6	Q9V4G6 drosophila
40	20	100.0	111	2 Q8RIS9	Q8RIS9 streptococc
41	20	100.0	111	16 Q8YKY3	Q8YKY3 anabaena sp
42	20	100.0	112	12 Q81667	Q81667 hepatitis c
43	20	100.0	113	9 Q38333	Q38333 lactococcus
44	20	100.0	115	11 Q64186	Q64186 cavia (guin
45	20	100.0	115	16 Q984W4	Q984W4 rhizobium 1

ALIGNMENTS

RESULT 1

P82952 ID P82952 PRELIMINARY; PRT; 25 AA.
AC P82952;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Seed allergenic protein (Fragment).
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. GENCO; TISSUE=SEED;
RA Poltronieri P., Cappello M.S., Dohmae N., Conti A., Zacheo G.;
RT "Identification of IgE-binding proteins as 2S albumin and conglutin in
RT almond (Prunus dulcis) seed.";
RL Submitted (FEB-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: MAY FUNCTION AS SEED STORAGE PROTEINS. HAS AN IGE-
CC BINDING ACTIVITY.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- TISSUE SPECIFICITY: SEED.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING SEED MATURATION.
KW Allergen; IgE-binding protein; Seed storage protein.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2872 MW; A04173DC65B48C99 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 21 DRAT 24

RESULT 2

```
Q9IKR6
ID Q9IKR6 PRELIMINARY; PRT; 27 AA.
AC Q9IKR6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
GN POL.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JBI;
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
Rameau J.L., Duffaut M., Durand D., Puel J.;
RT "Longitudinal analysis of Hepatitis C virus replication and liver
RT fibrosis progression in renal transplant recipients.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207348; AAF76589.1; -.
DR InterPro; IPR002531; HCV_NSI.
DR Pfam; PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2768 MW; 2F457A89D7D2263 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 10 DRAT 13

RESULT 3
Q8X254
ID Q8X254 PRELIMINARY; PRT; 31 AA.
AC Q8X254;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 3.6 kDa protein (Fragment).
OS Halobacterium volcanii (Haloflex volcanii).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2246;
RN [1]
RP SEQUENCE FROM N.A.
RA Bidle K.A.;
RT "Differential expression of genes influenced by changing salinity
RT using RNA arbitrarily primed PCR in the archaeal halophile, Haloflex
RT volcanii.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF454092; AAL57845.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 31
SQ SEQUENCE 31 AA; 3609 MW; 682119419273C9D0 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 27 DRAT 30

RESULT 4
Q8XXT7
ID Q8XXT7 PRELIMINARY; PRT; 48 AA.
AC Q8XXT7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC2026.
```

```
Q9URG8
ID Q9URG8 PRELIMINARY; PRT; 45 AA.
AC Q9URG8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Carboxypeptidase 3, Cp3 (Fragments).
OS Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucoraceae;
OC Mucor.
OX NCBI_TaxID=4841;
RN [1]
RP SEQUENCE.
RX MEDLINE=92105011; PubMed=1729237;
RA DiSanto M.E., Li Q.H., Logan D.A.;
RT "Purification and characterization of a developmentally regulated
RT carboxypeptidase from Mucor racemosus.";
RL J. Bacteriol. 174:447-455(1992).
FT NON_TER 1
FT NON_TER 28
FT NON_TER 45
FT NON_TER 45
SQ SEQUENCE 45 AA; 5103 MW; A9F3B7471465465A CRC64;

Query Match 100.0%; Score 20; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 16 DRAT 19

RESULT 5
Q9S604
ID Q9S604 PRELIMINARY; PRT; 46 AA.
AC Q9S604;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Sigma factor (fragment).
GN FCI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99242013; PubMed=10227474;
RA Reingold J., Starr N., Maurer J., Lee M.D.;
RT "Identification of a new Escherichia coli She haemolysin homolog in
RT avian E. coli.";
RL Vet. Microbiol. 66:125-134(1999).
DR EMBL; AF052225; AAD28082.1; -.
FT NON_TER 46
FT NON_TER 46
SQ SEQUENCE 46 AA; 5222 MW; 42D98F2A02BE4D19 CRC64;

Query Match 100.0%; Score 20; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 3 DRAT 6

RESULT 6
Q8XXT7
ID Q8XXT7 PRELIMINARY; PRT; 48 AA.
AC Q8XXT7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC2026.
```


GN RSC2026 OR RS03598.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN (1)_TaxID=305;
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choine N., Claudel-Renard C., Cunnac S., Denange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex I.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646067; CAD15728.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 48 AA; 5473 MW; 66CDA3D6402F3718 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
|||
Db 31 DRAT 34

RESULT 7
Q9LJS3
ID Q9LJS3 PRELIMINARY; PRT; 49 AA.
AC Q9LJS3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genomic DNA, chromosome 3, TAC clone: K11J14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
DR EMBL; AP000411; BAB02149.1; --
SQ SEQUENCE 49 AA; 5462 MW; 52A7EA47F98D5126 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
|||
Db 5 DRAT 8

RESULT 8
Q90XF3
ID Q90XF3 PRELIMINARY; PRT; 53 AA.
AC Q90XF3;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-binding protein H (Fragment).
OS Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN (1)
RP SEQUENCE FROM N.A.
RA Bader A.G., Schneider M.L., Bister K., Hartl M.;
RT "TO3, a target of the v-Jun transcription factor, encodes a protein
RT with transforming activity related to human microspherule protein 1
RT (MCRS1).";
RL Oncogene 0:0-0(2001).
DR EMBL; AF390035; AAK73733.1; --
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 6325 MW; 325C0B53BE2F8D27 CRC64;

Query Match 100.0%; Score 20; DB 13; Length 53;
Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
|||
Db 18 DRAT 21

RESULT 9
Q9U4L5
ID Q9U4L5 PRELIMINARY; PRT; 60 AA.
AC Q9U4L5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Mitotic cyclin-like protein (fragment).
GN CYC.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CU427;
RA Zhang H., Huang X., Berger J.D.;
RL "A cyclin-like sequence in Tetrahymena thermophila.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE CYCLIN FAMILY.
DR EMBL; AF165220; AAF15554.1; --
DR HSP; P30274; 1VIN.
DR InterPro; IPR004366; Cyclin.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cell cycle; Cell division; Cyclin.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7020 MW; 803EC056F39D45D2 CRC64;

Query Match 100.0%; Score 20; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
|||
Db 35 DRAT 38

RESULT 10
Q65240
ID Q65240 PRELIMINARY; PRT; 63 AA.

```

AC Q65240;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF j16L.
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarivirus.
OX NCBI_TaxID=10497;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RX MEDLINE=94014996; PubMed=8409937;
RA Vydelingum S., Baylis S.A., Bristow C., Smith G.L., Dixon L.K.;
RT "duplicated genes within the variable right end of the genome of a
pathogenic isolate of African swine fever virus."
RL J. Gen. Virol. 74:2125-2130(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RA Dixon L.K., Twigg S.R.F., Baylis S.A., Vydelingum S., Bristow C.,
HAMMOND J.M., Smith G.L.;
RT "Nucleotide sequence of a 55 kbp region from the right end of the
genome of a pathogenic African swine fever virus isolate (Malawi
LIL20/1)."
RL J. Gen. Virol. 7:1655-1694(1994).
DR EMBL: X71982; CAA50836.1; -; 7C1AC866CDDP4EEB CRC64;
SQ SEQUENCE 63 AA; 6851 MW; 7C1AC866CDDP4EEB CRC64;

Query Match 100.0%; Score 20; DB 12; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 36 DRAT 39

RESULT 11
Q9HQ75 PRELIMINARY; PRT; 63 AA.
AC Q9HQ75;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Vng1295h.
GN Vng1295h.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
Ezhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl Acad. Sci U.S.A. 97:12176-12181(2000).
DR EMBL: AS005052; AAC19642.1; -;
KW Complete proteome.
SQ SEQUENCE 63 AA; 6465 MW; 3A98B3C0FC089A68 CRC64;

Query Match 100.0%; Score 20; DB 17; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 36 DRAT 39

```

```

Db 30 DRAT 33

RESULT 12
Q9AYU9 PRELIMINARY; PRT; 64 AA.
AC Q9AYU9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 7.6 kDa protein (orf9).
GN ORF9.
OS Lactococcus lactis bacteriophage Tuc2009,
bacteriophage bIL285, and
bacteriophage bIL286.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=35241, 151535, 151536;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Lactococcus lactis bacteriophage Tuc2009;
RA van Sinderen D., van de Guchte M., Seegers J.F.M.L., Fitzgerald G.F.;
RT "Molecular analysis of the temperate lactococcal phage Tuc2009."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Phage bIL285, and Phage bIL286;
RX MEDLINE=21113149; PubMed=11160885;
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RT "Analysis of six prophages in Lactococcus lactis IL1403: different
genetic structure of temperate and virulent phage populations."
RL Nucleic Acids Res. 29:644-651(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Phage bIL285, and Phage bIL286;
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF109874; AAK19849.1; -;
DR EMBL: AF323668; AAK08234.1; -;
DR EMBL: AF323669; AAK08296.1; -;
KW Hypothetical protein.
SQ SEQUENCE 64 AA; 7573 MW; 740DC4287228266F CRC64;

Query Match 100.0%; Score 20; DB 9; Length 64;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 47 DRAT 50

RESULT 13
Q8W0J3 PRELIMINARY; PRT; 64 AA.
AC Q8W0J3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE P0007F06.9 protein.
GN P0007F06.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0007F06."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003223; BAB84370.1; -;

```

SQ SEQUENCE 64 AA; 7286 MW; 21256F1F01116535 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 64;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 48 DRAT 51

RESULT 14

Q9CFP1

ID Q9CFP1 PRELIMINARY; PRT; 64 AA.

AC Q9CFP1;

DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Prophage p13 protein 53.

GN P1353 OR L14426.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Lactococcus.

OX NCBI_TaxID=1360;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=IL1403;

RX MEDLINE=21235186; PubMed=11337471;

RA Holotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,

RA Weissenbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

lactis ssp. lactis IL1403."

RL Genome Res. 11:731-753(2001).

DR EMBL; AE06373; AAK05524.1; -.

KW Complete proteome.

SQ SEQUENCE 64 AA; 7573 MW; 740DC4287228266F CRC64;

Query Match

Best Local Similarity 100.0%; Score 20; DB 16; Length 64;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 47 DRAT 50

RESULT 15

Q9NTN6

ID Q9NTN6 PRELIMINARY; PRT; 66 AA.

AC Q9NTN6;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)

DE BB94P8.1 (K1A11326) (Fragment).

GN BB94P8.1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RA Wilson S.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL133274; CAB92740.1; -.

FT NON_TER 1

FT NON_TER 66

SQ SEQUENCE 66 AA; 7022 MW; 218B52004B6E6AA CRC64;

Query Match

Best Local Similarity 100.0%; Score 20; DB 4; Length 66;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 48 DRAT 51

Search completed: May 6, 2003, 15:02:48
Job time : 16.2 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:24 ; Search time 4 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: US-09-851-422B-2

Perfect score: 20
Sequence: 1 DRAT 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results' predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	67	1 YIFL_ECOLI	P39166 escherichia
2	20	100.0	70	1 PYRH_THETH	P43891 thermus the
3	20	100.0	77	1 FYAQ_VACCC	P20526 vaccinia vi
4	20	100.0	84	1 SCX7_TITBA	P56611 titiyus bali
5	20	100.0	84	1 SCX7_TITSE	P15226 titiyus serr
6	20	100.0	84	1 SCX7_TITST	P56612 titiyus stig
7	20	100.0	91	1 ACPA_RHET	P24901 rhizobium e
8	20	100.0	91	1 YA4D_SCHPO	Q09730 schizosacch
9	20	100.0	104	1 URE2_MYCTU	P50048 mycobacteri
10	20	100.0	109	1 VETM_NPVOP	O10307 orgyia pseu
11	20	100.0	121	1 RNPA_NEIMA	O9jw46 neisseria m
12	20	100.0	121	1 RNPA_NEIMB	O9jx56 neisseria m
13	20	100.0	125	1 AMPA_SYNP7	Q06885 synchococc
14	20	100.0	127	1 TRBC_RHFSN	P55396 rhizobium s
15	20	100.0	139	1 ADF3_MAIZE	P06714 zeamays (m
16	20	100.0	141	1 HBAT_HORSE	P43758 haemophilus
17	20	100.0	145	1 DKSA_HABIN	P57294 buchnera ap
18	20	100.0	151	1 DKSA_BUCAI	Q08950 gallus gall
19	20	100.0	160	1 MD1_CHICK	P41443 escherichia
20	20	100.0	169	1 GSPH_ECOLI	P23484 escherichia
21	20	100.0	173	1 PECI_ECOLI	O55024 synchococc
22	20	100.0	176	1 DP5A_SYNP7	O53717 mycobacteri
23	20	100.0	179	1 PYRE_MYCTU	Q9pzh6 ureaplasma
24	20	100.0	181	1 IPYR_UREPA	P40178 streptomyc
25	20	100.0	182	1 PURE_STRPR	Q9x8r7 streptomyc
26	20	100.0	182	1 PYRE_STRCO	P26427 escherichia
27	20	100.0	186	1 AHPC_ECOLI	P19479 salmonella
28	20	100.0	186	1 AHPC_ECOLI	P13076 halobacteri
29	20	100.0	193	1 FLA3_HALN1	P20771 rhodococc
30	20	100.0	199	1 YNHA_RHDER	O14807 homo sapien
31	20	100.0	208	1 RASM_HUMAN	O08989 mus musculu
32	20	100.0	208	1 RASM_MOUSE	P97538 rattus norv
33	20	100.0	208	1 RASM_RAT	

RESULT 1
YIFL_ECOLI STANDARD; PRT; 67 AA.
AC P39166; Q8X3Y5;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein YifL.
GN YIFL OR B3808.1 OR Z5325 OR ECS4737.1.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Glaser P., Sismeiro O., Danchin A.;
RI Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rued K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
bacterial genome.";

ALIGNMENTS

34 20 100.0 226 1 PROB_STRAG P09879 streptococc
35 20 100.0 231 1 FLGH_PSEPU Q52081 pseudomonas
36 20 100.0 236 1 PLL_SHEEP P16038 ovis aries
37 20 100.0 240 1 PYRH_AQUAE O69299 aquifex aeo
38 20 100.0 240 1 PYRH_BACSU O31749 bacillus su
39 20 100.0 243 1 TONE_PSEPU O05613 pseudomonas
40 20 100.0 253 1 CFAD_HUMAN P07146 homo sapien
41 20 100.0 253 1 YE24_MYCTU P22866 mycobacteri
42 20 100.0 254 1 GYLX_STRGR P22866 streptomyc
43 20 100.0 259 1 RPOD_HALMA Q00813 halocaula
44 20 100.0 260 1 PYRH_SVNY3 P74457 synchocyst
45 20 100.0 260 1 YH20_PASMU Q9cka5 pasteurella

RL Nucleic Acids Res. 22:4756-4767(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M87049; -; NOT ANNOTATED_CDS.
DR EMBL; AE000457; -; NOT ANNOTATED_CDS.
DR EMBL; X66782; -; NOT ANNOTATED_CDS.
DR EMBL; AE005612; AGS9001.1; -;
DR EMBL; AP002567; -; NOT ANNOTATED_CDS.
DR EcoGene; EG12353; yifL.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 60 67 DGSOVNY -> MVHPR (IN REF. 2).
SQ SEQUENCE 67 AA; 7177 MW; 4A52CF48A555A7F9 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 55 DRAT 58
|||||

RESULT 2
ID PYRH_TETH STANDARD; PRT; 70 AA.
AC P43891;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Uridylate kinase (EC 2.7.4.-) (UK) (Uridine monophosphate kinase)
DE (UMP kinase) (Fragment).
GN PYRH.

OS Thermus thermophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermacaeae; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=96184901; PubMed=8617268;
RA Blank J., Nock S., Kreutzer R., Sprinzl M.;
RT "Elongation factor Ts from Thermus thermophilus -- overproduction in
RT Escherichia coli, quaternary structure and interaction with
RT elongation factor Tu".
RL Eur. J. Biochem. 236:222-227(1996).
CC -|- FUNCTION: URIDINE MONOPHOSPHATE KINASE (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: ATP + UMP = ADP + UDP.
CC -|- ENZYME REGULATION: REGULATED BY GUANINE NUCLEOTIDES AND UTP
CC (BY SIMILARITY).
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -|- SIMILARITY: TO OTHER UMP KINASE; SOME, TO ASPARTOKINASES AND TO
CC GLUTAMATE KINASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X63598; CAAS58579.1; -;
DR InterPro: IPR001048; Aa-kinase.
DR Pfam: PF00696; aakinease; 1.
KW Transferase; Kinase; Pyrimidine biosynthesis.
FT NON_TER 70 70

SQ SEQUENCE 70 AA; 7629 MW; 7B8AB85DDA977CEB CRC64;

Query Match 100.0%; Score 20; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 64 DRAT 67
|||||

RESULT 3
ID YVAQ_VACC STANDARD; PRT; 77 AA.
AC P20526;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 8.6 kDa protein.
GN A ORF Q.

OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "The complete DNA sequence of vaccinia virus".
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.

RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'";
RL Virology 179:517-563(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M35037; AAA48178.1; -;
RX PIR; D42525; D42525.
KW Hypothetical protein.
SQ SEQUENCE 77 AA; 8635 MW; 12BBC61FF7472771 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 77;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 23 DRAT 26
|||||

RESULT 4
ID SCX7_TITBA STANDARD; PRT; 84 AA.
AC P56511;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Toxin gamma precursor.

OS Titus baliensis (Brazilian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Tityus.
OX NCBI_TaxID=50343;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-81.
 RC TISSUE=Venom;
 RX MEDLINE=96190713; PubMed=8611151;
 RA Becerril B., Corona M., Coronas F.I., Zamudio F.,
 RA Calderon-Aranda E.S., Fletcher P.L. Jr., Martin B.M., Possani L.D.;
 RT "Toxic peptides and genes encoding toxin gamma of the Brazilian
 RT scorpions *Tityus bahiensis* and *Tityus stigmurus*.";
 RL Biochem. J. 313:753-760(1996).
 CC -!- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
 CC OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
 CC BETA-TOXIN SUBFAMILY.
 DR HSSP; P01484; IPTX.
 DR InterPro; IPR003614; Knot1.
 DR InterPro; IPR002061; Scorpion_toxinL.
 DR Pfam; PF00537; toxin_3; 1.
 DR ProDom; PD000908; Scorpion_toxinL; 1.
 DR SMART; SM00505; Knot1; 1.
 KW Neurotoxin; Sodium channel inhibitor; Amidation; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 81 TOXIN GAMMA.
 FT DISULFID 31 81 BY SIMILARITY.
 FT DISULFID 35 57 BY SIMILARITY.
 FT DISULFID 43 62 BY SIMILARITY.
 FT DISULFID 47 64 BY SIMILARITY.
 FT MOD_RES 81 81 AMIDATION (G-82 PROVIDE AMIDE GROUP)
 FT (PROBABLE).
 SQ SEQUENCE 84 AA; 9384 MW; A24A2ACAF768136 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 84;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRAT 4
 ||||
 Db 75 DRAT 78

RESULT 5
 SCX7 TITSE STANDARD; PRT; 84 AA.
 AC F12226;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Toxin VII precursor (TstX-VII) (Tityustoxin VII) (Ts VII) (Toxin
 DE II-1) (Toxin III-10) (Toxin gamma) (Toxin T2-IV).
 OS Tityus serrulatus (Brazilian scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Butthidea; Butthidea; Tityus.
 OX NCBI_TaxID=6887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom;
 RX MEDLINE=92290001; PubMed=1339357;
 RA Martin-Baucalire M.-F., Ceard B., Ribeiro A.M., Diniz C.R., Rochat H.,
 RA Bougis P.E.;
 RT "Molecular cloning and nucleotide sequence analysis of a cDNA
 RT encoding the main beta-neurotoxin from the venom of the South
 RT American scorpion *Tityus serrulatus*.";
 RL FEBS Lett. 302:220-222(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom;
 RX MEDLINE=94063080; PubMed=8243666;
 RA Becerril B., Corona M., Mejia M.C., Martin B.M., Bolivar F.,
 RA Possani L.D.;
 RT "The genomic region encoding toxin gamma from the scorpion *Tityus*
 RT serrulatus contains an intron.";
 RL FEBS Lett. 335:6-8(1993).
 RN [3]
 RP SEQUENCE OF 21-81.

RC TISSUE=Venom;
 RX MEDLINE=84307524; PubMed=6477555;
 RA Bechis G., Sampieri F., Yuan P.-M., Brando T., Martin M.-F.,
 RA Diniz C.R., Rochat H.;
 RT "Amino acid sequence of toxin VII, a beta-toxin from the venom of the
 RT scorpion *Tityus serrulatus*.";
 RL Biochem. Biophys. Res. Commun. 122:1146-1153(1984).
 RN [4]
 RP SEQUENCE OF 21-81.
 RC TISSUE=Venom;
 RX MEDLINE=91131623; PubMed=1993690;
 RA Possani L.D., Martin B.M., Fletcher M.D., Fletcher P.L. Jr.;
 RT "Discharge effect on pancreatic exocrine secretion produced by toxins
 RT purified from *Tityus serrulatus* scorpion venom.";
 RL J. Biol. Chem. 266:3178-3185(1991).
 RN [5]
 RP SEQUENCE OF 21-62.
 RC TISSUE=Venom;
 RX Possani L.D., Martin B.M., Mochca-Morales J., Svendsen I.;
 RA "Purification and chemical characterization of the major toxins from
 RA the venom of the Brazilian scorpion *Tityus serrulatus* Lutz and Mello.";
 RL Carlsberg Res. Commun. 46:195-205(1981).
 RN [6]
 RP SEQUENCE OF 21-46.
 RC TISSUE=Venom;
 RX MEDLINE=92023285; PubMed=1926167;
 RA Sampaio S.V., Arantes E.C., Ricardo Neto F.,
 RA Giglio J.R.;
 RT "Further characterization of toxins TIIV (TstX-III) and T2IV from
 RT *Tityus serrulatus* scorpion venom.";
 RL Toxicon 29:663-672(1991).
 CC -!- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
 CC OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
 CC LD(50) IS 0.6 NG. INTRACEREBROVENTRICULAR.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
 CC BETA-TOXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X66256; CAA46982.1; -.
 DR EMBL; S66941; AAB29128.1; -.
 DR PIR; A39510; A39510.
 DR PIR; S21158; S21158.
 DR PIR; S39438; S39438.
 DR HSSP; P01484; IPTX.
 DR InterPro; IPR003614; Knot1.
 DR InterPro; IPR002061; Scorpion_toxinL.
 DR Pfam; PF00537; toxin_3; 1.
 DR ProDom; PD000908; Scorpion_toxinL; 1.
 DR SMART; SM00505; Knot1; 1.
 KW Neurotoxin; Sodium channel inhibitor; Amidation; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 81 TOXIN VII.
 FT DISULFID 31 81 BY SIMILARITY.
 FT DISULFID 35 57 BY SIMILARITY.
 FT DISULFID 43 62 BY SIMILARITY.
 FT DISULFID 47 64 BY SIMILARITY.
 FT MOD_RES 81 81 AMIDATION (G-82 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 84 AA; 9382 MW; 430F3FCA74823E77 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 84;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRAT 4
 ||||

```

Db 75 DRAT 78

RESULT 6
SCX7_TITST STANDARD; PRT; 84 AA.
AC P56612;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Toxin gamma precursor.
OS Tityus stigmurus (Brazilian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butioidae; Tityidae; Tityus.
OX NCBI_TaxID=50344;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-81 FROM N.A.
RC TISSUE=Venom;
RX MEDLINE=96190713; PubMed=8611151;
RA Becerril B., Corona M., Coronas F.I., Zamudio F.,
Caldaron-Aranda E.S., Fletcher P.L. Jr., Martin B.M., Poseani L.D.;
"toxic peptides and genes encoding toxin gamma of the Brazilian
scorpions Tityus bahiensis and Tityus stigmurus.";
RL Biochem. J. 313:753-760(1996).
CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC BETA-TOXIN SUBFAMILY.
DR HSPSP; P01484; IPTX.
DR InterPro; IPR003614; Kntcl.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; toxin_3; 1.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Kntcl; 1.
DR Neurotoxin; Sodium channel inhibitor; Amidation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 81 TOXIN GAMMA.
FT DISULFID 31 81 BY SIMILARITY.
FT DISULFID 35 57 BY SIMILARITY.
FT DISULFID 43 62 BY SIMILARITY.
FT DISULFID 47 64 BY SIMILARITY.
FT MOD_RES 81 81 AMIDATION (G-82 PROVIDE AMIDE GROUP)
(PROBABLE).
FT SEQUENCE 84 AA; 9366 MW; 460653ABAE1F7877 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 75 DRAT 78

RESULT 7
ACPA_RHIET STANDARD; PRT; 91 AA.
ID ACPA_RHIET
AC P24901; O88153;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein acpXL (ORF*).
GN ACPXL.
OS Rhizobium etli, and
Rhizobium leguminosarum (biovar viciae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=29449, 387;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=R.etli; STRAIN=CE3;
RX PubMed=11717256;

Lopez O., Morera C., Miranda-Rios J., Girard L., Romero D.,
Soberon M.;
"Regulation of gene expression in response to oxygen in Rhizobium
etli: role of FnrN in fixNOQP expression and in symbiotic nitrogen
fixation.";
J. Bacteriol. 183:6999-7006(2001).
RN [2]
RP SEQUENCE FROM N.A., SEQUENCE OF 1-27; 29-40 AND 68-74,
CHARACTERIZATION, AND MASS SPECTROMETRY.
RC SPECIES=R.etli; STRAIN=VF39;
RX MEDLINE=97113013; PubMed=8943266;
RA Broek K.A., Carlson R.W., Raetz C.R.H.;
"A special acyl carrier protein for transferring long hydroxylated
fatty acids to lipid A in Rhizobium.";
J. Biol. Chem. 271:32126-32136(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=R.etli; STRAIN=VF39;
RX MEDLINE=91080854; PubMed=2175385;
RA Colonna-Romano S., Arnold W., Schlueter A., Boistard P., Puehler A.,
Priefer U.B.;
"An Fnr-like protein encoded in Rhizobium leguminosarum biovar viciae
shows structural and functional homology to Rhizobium meliloti
FixK.";
Mol. Gen. Genet. 223:138-147(1990).
RN [4]
RP SEQUENCE OF 1-50 FROM N.A.
RC SPECIES=R.etli; STRAIN=VF39;
RX MEDLINE=91080854; PubMed=2175385;
RA Colonna-Romano S., Arnold W., Schlueter A., Boistard P., Puehler A.,
Priefer U.B.;
"An Fnr-like protein encoded in Rhizobium leguminosarum biovar viciae
shows structural and functional homology to Rhizobium meliloti
FixK.";
Mol. Gen. Genet. 223:138-147(1990).
CC -1- FUNCTION: Carrier of the growing fatty acid chain in fatty acid
biosynthesis. Is involved in the transfer of long hydroxylated
fatty acids to lipid A. Is acylated predominantly with 27-
hydroxyoctacosanoic acid.
CC -1- PATHWAY: Lipid A biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
serine of apo-ACP by acps. This modification is essential for
activity because fatty acids are bound in thioester linkage to the
sulphydryl of the prosthetic group.
CC -1- MASS SPECTROMETRY: MW=10486.0; METHOD=Electrospray; RANGE=Isoform
derivatized with 4'-phosphopantetheine.
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; AF083916; AAC34462.1;
CC EMBL; AF081835; AAC32203.1;
CC EMBL; AF510733; AAM44294.1; ALT_INIT.
CC EMBL; X55788; CAA39311.1; ALT_INIT.
CC PIR; S11950; S11950.
CC InterPro; IPR003231; Acyl carrier.
CC InterPro; IPR003880; Ppantne_attach.
CC Pfam; PF00550; pp-binding; 1.
CC ProDom; PD000887; Acyl carrier; 1.
CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
CC PROSITE; PS00075; ACP DOMAIN; 1.
KW Lipid synthesis; Lipid A biosynthesis; Fatty acid biosynthesis;
KW Phosphopantetheine.
FT INIT_MET 0 0
FT BINDING 36 36 PHOSPHOPANTETHEINE.
SQ SEQUENCE 91 AA; 10146 MW; 8E2E71DBAAB7B5A3 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 91;

```


Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 18 DRAT 21

RESULT 8

Y44D_SCHPO STANDARD; PRT; 91 AA.

AC Q09730;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C31A2.13c in chromosome 1.
GN SPAC31A2.13C
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lynne R., Stewart A.,
RA Spouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grynoprez B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Mostl D., Hilbert H.,
RA Borzyski K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
RA Dominguez A., Revuelta J.B., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z50113; CAA90471.1; -

CC KW Hypothetical protein; Transmembrane.

CC FT TRANSMEM 68 84 POTENTIAL.

CC SQ SEQUENCE 91 AA; 10542 MW; 78F695CE8B16CCB CRC64;

Query Match 100.0%; Score 20; DB 1; Length 91;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 37 DRAT 40

RESULT 9

URE2_MYCTU STANDARD; PRT; 104 AA.
AC P50048;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN UREB OR RV1849 OR MT1897 OR MTCY359.24C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H378V;
RX MEDLINE=96004620; PubMed=7568014;
RA Keyrat J.M., Berthet F.X., Gicquel B.;
RT "The urease locus of Mycobacterium tuberculosis and its utilization
RT for the demonstration of allelic exchange in Mycobacterium bovis
RT bacillus Calmette-Guerin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:8768-8772(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ersmann;
RX MEDLINE=96032403; PubMed=7559354;
RA Clemens D.L., Lee B.Y., Horwitz M.A.;
RT "Purification, characterization, and genetic analysis of
RT Mycobacterium tuberculosis urease, a potentially critical determinant
RT of host-pathogen interaction.";
RL J. Bacteriol. 177:5644-5652(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H378V;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Rutter S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Sulston J.E., Taylor K., Whitehead S., Squares R.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleisachmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- SUBUNIT: (ALPHA, BETA, GAMMA)(3) (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER UREASES BETA SUBUNITS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; L41141; AAC37006.1; -
CC EMBL; U33011; AAC43474.1; -
DR
DR

```

DR EMBL; Z83859; CAB06138.1; -.
DR EMBL; AB007047; AAK46168.1; -.
DR HSSP; PI8315; 1FWB.
DR TIGR; MT1897; -.
DR TubercuList; Rv1849; -.
DR InterPro; IPR002019; Urease_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 104 AA; 11190 MW; D621CE43A47304E0 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 53 DRAT 56

RESULT 10
VETM NPVOP STANDARD; PRT; 109 AA.
AC O10307;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ECORI-T site protein ETM homolog.
GN ETM.
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMPNV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H.; Russell R.R.; Funk C.J.; Evans J.; Harwood S.;
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome."
RL Virology 229:381-399(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U75930; AAC59051.1; -.
DR Early protein.
KW SEQUENCE 109 AA; 12053 MW; 597E7F7137389454 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 33 DRAT 36

RESULT 11
RNPA NEIMA STANDARD; PRT; 121 AA.
AC Q9JW46;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease P protein component (EC 3.1.26.5) (RNaseP protein) (RNase
DE P protein) (Protein C5).
GN RNPA OR NMA0550.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;

```

```

GN RNPA OR NMA0550.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J.; Achtman M.; James K.D.; Bentley S.D.; Churcher C.;
RA Klee S.R.; Morelli G.; Basham D.; Brown D.; Chillingworth T.;
RA Davies R.M.; Davis P.; Devlin K.; Feltwell T.; Hamlin N.; Holroyd S.;
RA Jagels K.; Leather S.; Moule S.; Mungall K.; Quail M.A.;
RA Rajandream M.A.; Rutherford K.M.; Simmonds M.; Skelton J.;
RA Whitehead S.; Spratt B.G.; Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -----
CC -I- FUNCTION: RNaseP catalyzes the removal of the 5'-leader sequence
CC from pre-tRNA to produce the mature 5'terminus. It can also
CC cleave other RNA substrates such as 4.5S RNA. The protein
CC component plays an auxiliary but essential role in vivo by binding
CC to the 5'-leader sequence and broadening the substrate specificity
CC of the ribozyme (By similarity).
CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -I- SUBUNIT: Consists of a catalytic RNA component (M1 or rnpB) and a
CC protein subunit (By similarity).
CC -I- SIMILARITY: BELONGS TO THE RNPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL162753; CAB83843.1; -.
DR InterPro; IPR001000; Ribonuclease_P.
DR Pfam; PF00825; Ribonuclease_P; 1.
DR ProDom; PD003629; Ribonuclease_P; 1.
DR TIGRFAMs; TIGR00188; rnpA; 1.
DR PROSITE; PS00648; RIBONUCLEASE_P; FALSE NEG.
KW Hydrolase; Nuclease; Endonuclease; tRNA processing; RNA-binding;
KW Complete proteome.
SQ SEQUENCE 121 AA; 14288 MW; 860703F33865E0F2 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 94 DRAT 97

RESULT 12
RNPA NEIMB STANDARD; PRT; 121 AA.
AC Q9JXS6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease P protein component (EC 3.1.26.5) (RNaseP protein) (RNase
DE P protein) (Protein C5).
GN RNPA OR NMB1905.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;

```

RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Rickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: RNaseP catalyzes the removal of the 5'-leader sequence
CC from pre-tRNA to produce the mature 5'terminus. It can also
CC cleave other RNA substrates such as 4.5S RNA. The protein
CC component plays an auxiliary but essential role in vivo by binding
CC to the 5'-leader sequence and broadening the substrate specificity
CC of the ribozyme (By similarity).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -!- SUBUNIT: Consists of a catalytic RNA component (M1 or rnpB) and a
CC protein subunit (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RNPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002540; AAF42235.1; -;
CC TIGR; NMB1905; -;
CC InterPro; IPR000100; Ribonuclease P.
CC Pfam; PF00825; Ribonuclease_P; 1.
CC ProDom; PD003629; Ribonuclease_P; 1.
CC TIGRFAMs; TIGR00188; rnpA; 1.
CC PROSITE; PS00648; RIBONUCLEASE_P; FALSE NEG.
CC HydroLase; Nuclease; Endonuclease; tRNA processing; RNA-binding;
CC Complete proteome.
CC SEQUENCE 121 AA; 14211 MW; D4DE8C0EA865E325 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 90;
Matches *4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 94 DRAT 97

RESULT 13
AMPA_SYN7
ID AMPA_SYN7 STANDARD; PRT; 125 AA.
AC O06865;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE [LAP] (Leucyl aminopeptidase) (Fragment).
GN PEPA OR LAP.
OS *Synechococcus* sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubio L.M., Flores E., Herrero A.;
RT "The narC locus of *Synechococcus* sp. strain PCC 7942 corresponds to a
RT mba gene for molybdopterine guanine dinucleotide biosynthesis.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
CC TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
CC UNSUBSTITUTED N-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES (BY

CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC including pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: MANGANESE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y13330; CAA73771.1; -;
CC HSPF; P00727; ILAM.
CC MEROPS; M17.002; -;
CC InterPro; IPR000819; Peptidase M17.
CC Pfam; PF00883; Peptidase M17; 1.
CC KW Hydrolase; Aminopeptidase; Manganese.
CC FT NON_TER 1
CC SEQUENCE 125 AA; 13409 MW; 42ABBD917A0AA971 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 2 DRAT 5

RESULT 14
TRBC_RHISN
ID TRBC_RHISN STANDARD; PRT; 127 AA.
AC P53396;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable conjugal transfer protein trbc.
GN TRBC OR Y4CN.
OS *Rhizobium* sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; *Rhizobium*.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between *Rhizobium* and legumes.";
RL Nature 387:394-401(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: STRONG, TO A.TUMEFACIENS TI PLASMID TRBC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000068; AAB92429.1; -;
CC Conjugation; Plasmid; Transmembrane.
CC TRANSMEM 7 27 POTENTIAL.
CC TRANSMEM 50 70 POTENTIAL.
CC TRANSMEM 79 99 POTENTIAL.
CC SEQUENCE 127 AA; 12738 MW; 92D46CD03AA00DE4 CRC64;
CC

Query Match 100.0%; Score 20; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 111 DRAT 114

Db 48 DRAT 51

Search completed: May 6, 2003, 15:01:29
Job time : 6 secs

RESULT 15
ADF3_MAIZE
ID ADF3_MAIZE STANDARD; PRT; 139 AA.
AC Q41764;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Actin-depolymerizing factor 3 (ADF 3) (ZmADFP3) (ZmADFP3).
GN ADF3 OR ADF3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_taxid=4577;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Leaf;
RX MEDLINE=96293540; PubMed=8691008;
RA Lopez I., Anthony R.G., Maciver S.K., Jiang C.J., Khan S., Weeds A.G.,
Hussey P.J.;
RT "Pollen specific expression of maize genes encoding actin
depolymerizing factor-like proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 93:7415-7420(1996).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98079240; PubMed=9418045;
RA Jiang C.J., Weeds A.G., Hussey P.J.;
RT "The maize actin-depolymerizing factor, ZmADFP3, redistributes to the
growing tip of elongating root hairs and can be induced to
translocate into the nucleus with actin."
RL Plant J. 12:1035-1043(1997).
CC -!- FUNCTION: ACTIN-DEPOLYMERIZING PROTEIN. SEVERES ACTIN FILAMENTS
(F-ACTIN) AND BINDS TO ACTIN MONOMERS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. As root hairs emerges and the
microfilament bundles redirect to the outgrowth ADF3 concentrates
at the tip of the emerging hair and remains in this position as
elongation proceeds.
CC -!- TISSUE SPECIFICITY: Expressed in all tissues except pollen.
CC -!- SIMILARITY: BELONGS TO THE ACTIN-BINDING PROTEINS ADF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97726; CAA66311.1; --
DR HSSP; Q39250; 1F7S.
DR InterPro; IPR002108; Actbind_cofln.
DR Pfam; PF00241; coflin_ADF; 1.
DR ProDom; PD002129; Actbind_cofln; 1.
DR SMART; SM00102; ADF; 1.
DR PROSITE; PS00325; ACTIN_DEPOLYMERIZING; 1.
KW Actin-binding; Multigene family.
FT DOMAIN 92 111 ACTIN-BINDING (POTENTIAL).
SQ SEQUENCE 139 AA; 15900 MW; 28FC8C518022C87 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:05 ; Search time 7 Seconds
(without alignments)
54.934 Million cell updates/sec

Title: US-09-851-422B-2

Perfect score: 20
Sequence: 1 DRAT 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	28	A44923	carboxypeptidase 3
2	20	100.0	63	F84284	hypothetical prote
3	20	100.0	64	S11950	hypothetical prote
4	20	100.0	64	B86803	prophage p13 prote
5	20	100.0	67	E86067	hypothetical prote
6	20	100.0	67	AD0919	probable lipoprote
7	20	100.0	70	S51096	UMP kinase - Therm
8	20	100.0	70	A60912	vitamin D receptor
9	20	100.0	77	D42525	A-ORF-Q protein -
10	20	100.0	84	S62867	toxin gamma precu
11	20	100.0	84	S62868	toxin gamma precu
12	20	100.0	84	S21158	neurotoxin TsVII p
13	20	100.0	87	B82666	hypothetical prote
14	20	100.0	91	S59647	probable SNARE pro
15	20	100.0	91	C83804	hypothetical prote
16	20	100.0	104	A70865	probable ureB prot
17	20	100.0	105	JQ2140	hypothetical 12.2K
18	20	100.0	105	JQ1823	SalFb protein - va
19	20	100.0	109	T10321	hypothetical prote
20	20	100.0	111	AC2497	hypothetical prote
21	20	100.0	113	AD2868	hypothetical prote
22	20	100.0	113	T13270	hypothetical prote
23	20	100.0	115	S68316	iron regulatory pr
24	20	100.0	118	E30540	Ig heavy chain V r
25	20	100.0	118	H97644	hypothetical prote
26	20	100.0	119	T11049	hypothetical prote
27	20	100.0	120	I39623	probable adenyate
28	20	100.0	121	F81973	probable ribonucle
29	20	100.0	121	H81027	ribonuclease P pro

30 20 100.0 122 2 AH2170
31 20 100.0 123 2 T07886
32 20 100.0 128 2 S40345
33 20 100.0 129 2 C83418
34 20 100.0 130 2 E87024
35 20 100.0 135 2 C82438
36 20 100.0 138 2 F81243
37 20 100.0 138 2 G82015
38 20 100.0 139 2 T02914
39 20 100.0 141 2 S72386
40 20 100.0 142 4 S07342
41 20 100.0 144 2 E72647
42 20 100.0 144 2 T01476
43 20 100.0 145 2 A64046
44 20 100.0 145 2 C82513
45 20 100.0 151 2 C84953

ferredoxin [import
protein kinase (EC
Ig kappa chain V-J
hypothetical prote
probable membrane
regulator of nucle
Dnak suppressor pr
probable Dnak supp
actin-depolymerizi
hypothetical prote
hypothetical hemog
hypothetical prote
dosage-dependent d
conserved hypothet
dnak suppressor pr

ALIGNMENTS

RESULT 1

A44923

carboxypeptidase 3 - Rhizomucor circinelloides f. lusitanicus (fragment)

C;Species: Rhizomucor circinelloides f. lusitanicus

C;Date: 01-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 04-Dec-1994

C;Accession: A44923

R;DiSanto, M.E.; Li, Q.H.; Logan, D.A.

J;Bacteriol. 174, 447-455, 1992

A;Title: Purification and characterization of a developmentally regulated carboxypeptid

A;Reference number: A44923; MUID:92105011; PMID:1729237

A;Accession: A44923

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-28 <DS>

A;Note: sequence extracted from NCBI backbone (NCBIP:75614)

Query Match 100.0%; Score 20; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 16 DRAT 19

RESULT 2

F84284

hypothetical protein Vng1295h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: F84284

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,

; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L

A;Title: Genome sequence of Halobacterium species NRC-1

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: F84284

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-63 <STO>

A;Cross-references: GB:AE004437; NID:G10580814; PIDN:ANG19642.1; GSPDB:GN00138

A;Gene: VNG1295H

C;Genetics:

QY 1 DRAT 4

Db 16 DRAT 19

Query Match 100.0%; Score 20; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 DRAT 33

RESULT 3

S11950
 hypothetical protein - Rhizobium leguminosarum (fragment)
 C:Species: Rhizobium leguminosarum
 C>Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 21-Jul-2000
 C:Accession: S11950
 R:Colonna-Romano, S.; Arnold, W.; Schlueter, A.; Boistard, P.; Puehler, A.; Priefer, U.E.
 Mol. Gen. Genet. 223, 138-147, 1990
 A>Title: An fmr-like protein encoded in Rhizobium leguminosarum biovar viciae shows stru
 A:Reference number: S11950; MUID:91080854; PMID:2175385
 A:Accession: S11950
 A:Molecule type: DNA
 A:Residues: 1-64 <COL>
 A:Cross-references: EMBL:X55789; NID:g311223; PIDN:CAA39311.1; PID:g3980213
 C:Superfamily: acyl carrier protein; acyl carrier protein homology
 C:Keywords: carrier protein
 F:15-64/Domain: acyl carrier protein homology (fragment) <ACP>

Query Match 100.0%; Score 20; DB 2; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 ||||
 Db 32 DRAT 35

RESULT 4

B86803
 prophage pi3 protein 53 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: B86803
 R:Boletini, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A:Reference number: A86825; MUID:21235186; PMID:11337471
 A:Accession: B86803
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-64 <STO>
 A:Cross-references: GB:AE005176; PID:g12724415; PIDN:NAK05524.1; GSPDB:GNO0146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: pi353

Query Match 100.0%; Score 20; DB 2; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 ||||
 Db 47 DRAT 50

RESULT 5

E86067
 hypothetical protein Z5325 [imported] - Escherichia coli (strain O157:H7, substrain EDL5
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: E86067
 R:Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: E86067
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-67 <STO>

A:Cross-references: GB:AE005174; NID:g12518676; PIDN:AAG59001.1; GSPDB:GNO0145; UWGP:Z5
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z5325

Query Match 100.0%; Score 20; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 ||||
 Db 55 DRAT 58

RESULT 6

AD0919
 probable lipoprotein yifL [imported] - Salmonella enterica subsp. enterica serovar Typhi
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A>Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AD0919
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AD0919
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-67 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD09374.1; PID:g16504492; GSPDB:GNO0176
 C:Genetics:
 A:Gene: yifL

Query Match 100.0%; Score 20; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 ||||
 Db 55 DRAT 58

RESULT 7

S51096
 UMP kinase - Thermus aquaticus (fragment)
 C:Species: Thermus aquaticus
 C>Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
 C:Accession: S51096
 R:Blank, J.; Grillenbeck, N.; Kreutzer, R.; Sprinzl, M.
 submitted to the EMBL Data Library, December 1994
 A:Description: Overproduction of Thermus thermophilus elongation factors G, Tu and Ts a
 A:Reference number: S51094
 A:Accession: S51096
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-70 <BLA>
 A:Cross-references: EMBL:X83598; NID:g619568; PIDN:CAA58579.1; PID:g619571
 A>Note: the source is designated as Thermus thermophilus
 C:Superfamily: uridine 5'-monophosphate kinase

Query Match 100.0%; Score 20; DB 2; Length 70;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 ||||
 Db 64 DRAT 67

RESULT 8

A60912
vitamin D receptor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Sep-1997
C:Accession: A60912
R:McDonnell, D.P.; Mangelsdorf, D.J.; Pike, J.W.; Haussler, M.R.; O'Malley, B.W.
Science 235, 1214-1217, 1987
A:Title: Molecular cloning of complementary DNA encoding the avian receptor for vitamin
A:Reference number: A60912; MUID:87149040; PMID:3029866
A:Accession: A60912
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-70 <MCD>
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; zinc finger
F:1-70/Domain: erba transforming protein homology (fragment) <ERBA>

Query Match 100.0%; Score 20; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||||
Db 8 DRAT 11

RESULT 9
D42525
A-ORF-Q protein - vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus
C:Note: host Homo sapiens (man)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 08-Apr-1994
C:Accession: D42525
R:Johnson, G.P.
submitted to GenBank, June 1990
A:Reference number: A33172
A:Accession: D42525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <JOH>

Query Match 100.0%; Score 20; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||||
Db 23 DRAT 26

RESULT 10
S62867
toxin gamma precursor - Tityus stigmurus
C:Species: Tityus stigmurus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S62867; S62865
R:Becerril, B.; Corona, M.; Coronas, F.I.V.; Zamudio, F.; Calderon-Aranda, E.S.; Fletcher
Biochem. J. 313, 753-760, 1996
A:Title: Toxic peptides and genes encoding toxin gamma of the Brazilian scorpions Tityus
A:Reference number: S62861; MUID:96190713; PMID:8611151
A:Accession: S62867
A:Molecule type: DNA
A:Residues: 1-84 <BBC>
A:Accession: S62865
A:Molecule type: protein
A:Residues: 20-81 <BEW>
C:Superfamily: scorpion neurotoxin
F:1-20/Domain: amidated carboxyl end; neurotoxin; venom
F:21-82/Product: toxin gamma #status predicted <SIG>
F:31-81,35-57,43-62,47-64/Disulfide bonds: #status predicted <MAT>
F:81/Modified site: amidated carboxyl end (Cys) (amide in mature form from following gly

Query Match 100.0%; Score 20; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||||
Db 75 DRAT 78

RESULT 11
S62868
toxin gamma precursor - Tityus bahiensis
C:Species: Tityus bahiensis
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S62868; S62861
R:Becerril, B.; Corona, M.; Coronas, F.I.V.; Zamudio, F.; Calderon-Aranda, E.S.; Fletcher
Biochem. J. 313, 753-760, 1996
A:Title: Toxic peptides and genes encoding toxin gamma of the Brazilian scorpions Tityus
A:Reference number: S62861; MUID:96190713; PMID:8611151
A:Accession: S62868
A:Molecule type: DNA
A:Residues: 1-84 <BEC>
A:Accession: S62861
A:Molecule type: protein
A:Residues: 21-81 <BEW>
C:Superfamily: scorpion neurotoxin
F:1-20/Domain: amidated carboxyl end; neurotoxin; venom
F:21-82/Product: toxin gamma #status experimental <MAT>
F:31-81,35-57,43-62,47-64/Disulfide bonds: #status predicted
F:81/Modified site: amidated carboxyl end (Cys) (amide in mature form from following gly

Query Match 100.0%; Score 20; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||||
Db 75 DRAT 78

RESULT 12
S21158
neurotoxin TsvII precursor - Brazilian scorpion
N:Alternate names: beta-neurotoxin; neurotoxin II-11; neurotoxin III-10; toxin gamma
C:Species: Tityus serrulatus (Brazilian scorpion)
C>Date: 22-Nov-1993 #sequence_revision 04-Oct-1996 #text_change 16-Jul-1999
C:Accession: S21158; S39438; A39510; S32790
R:Martin-Eauclaire, M.F.; Ceard, B.; Ribeiro, A.M.; Diniz, C.R.; Rochat, H.; Bougis, P.
FEBS Lett. 302, 220-222, 1992
A:Title: Molecular cloning and nucleotide sequence analysis of a cDNA encoding the main
A:Reference number: S21158; MUID:92290001; PMID:1339357
A:Accession: S21158
A:Molecule type: mRNA
A:Residues: 1-84 <MAR>
A:Cross-references: EMBL:X66256; NID:g312021; PIDN:CAA46982.1; PID:g312022
R:Becerril, B.; Corona, M.; Mejia, M.C.; Martin, B.M.; Lucas, S.; Bolivar, F.; Possani,
FEBS Lett. 335, 6-8, 1993
A:Title: The genomic region encoding toxin gamma from the scorpion Tityus serrulatus co
A:Reference number: S39438; MUID:94063080; PMID:8243666
A:Accession: S39438
A:Molecule type: DNA
A:Residues: 1-84 <BEC>
A:Cross-references: EMBL:S66941; NID:g453105; PIDN:AAB29128.1; PID:g453106
R:Possani, L.D.; Martin, B.M.; Fletcher, M.D.; Fletcher Jr., P.L.
J. Biol. Chem. 266, 3178-3185, 1991
A:Title: Discharge effect on pancreatic exocrine secretion produced by toxins purified
A:Reference number: A39510; MUID:91131623; PMID:1993690
A:Accession: A39510
A:Molecule type: protein
A:Residues: 21-81 <POS>
R:Possani, L.D.; Martin, B.M.; Svendsen, I.; Rode, G.S.; Erickson, B.W.
Biochem. J. 229, 739-750, 1985

A;Title: Scorpion toxins from Centruroides noxius and Tityus serrulatus. Primary structure
 A;Reference number: S32789; MUID:86025386; PMID:4052021

A;Accession: S32790
 A;Molecule type: protein
 A;Residues: 21-81 <P02>

C;Genetics:

A;Introns: 15/1

C;Superfamily: scorpion neurotoxin

C;Keywords: amidated carboxyl end; neurotoxin; venom

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-81/Product: neurotoxin TVII #status experimental <MAT>

F;31-81,35-57,43-62,47-64/Disulfide bonds: #status predicted

F;81/Modified site: amidated carboxyl end (Cys) (amide in mature form from following gly

Query Match 100.0%; Score 20; DB 2; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 75 DRAT 78

|||||

RESULT 13

B82666

hypothetical protein XP151 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C;Accession: B82666

R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: B82666

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-87 <SIM>

A;Cross-references: GB:AE003985; GB:AE003849; NID:g9106593; PIDN:AAF84360.1; GSPDB:GN001

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.A.; Canargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XP151

Query Match 100.0%; Score 20; DB 2; Length 87;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 82 DRAT 85

|||||

RESULT 14

S59647

probable SNARE protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 14-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000

C;Accession: T38611; S59647

R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z21731

A;Accession: T38611

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-91 <DE2>

A;Cross-references: EMBL:Z50113; NID:g914878; PIDN:CAA90471.1; PID:g914891; GSPDB:GN000

A;Experimental source: strain 972h-; Cosmid c31A2

C;Genetics:

A;Gene: SPAC31A2.13c

A;Map position: 1

A;Introns: 39/3

Query Match 100.0%; Score 20; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 37 DRAT 40

|||||

RESULT 15

C83804

hypothetical protein BH1235 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: C83804

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: C83804

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-91 <STO>

A;Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04954.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1235

Query Match 100.0%; Score 20; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 31 DRAT 34

|||||

Search completed: May 6, 2003, 15:03:29

Job time : 9 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 15:01:10 ; Search time 8 Seconds
(without alignments)
43.143 Million cell updates/sec

Title: US-09-851-422b-2

Perfect score: 20

Sequence: 1 DRAT 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB.pap.*
2: /cgn2_6/ptodata/1/pubpaa/PT NEW PUB.pap.*
3: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB.pap.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB.pap.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/1/pubpaa/US09 NEW PUB.pap.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB.pap.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/1/pubpaa/US60 NEW PUB.pap.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	19	9 US-10-067-813-5	Sequence 5, Appl
2	20	100.0	26	9 US-10-043-487-532	Sequence 532, App
3	20	100.0	50	9 US-10-142-231-37	Sequence 37, Appl
4	20	100.0	57	10 US-09-864-761-37634	Sequence 37634, A
5	20	100.0	66	10 US-09-864-761-36163	Sequence 36163, A
6	20	100.0	72	9 US-10-102-806-807	Sequence 807, App
7	20	100.0	72	10 US-09-864-761-34962	Sequence 34962, A
8	20	100.0	78	10 US-09-864-761-34713	Sequence 34713, A
9	20	100.0	84	9 US-10-092-154-620	Sequence 620, App
10	20	100.0	84	10 US-09-764-847-620	Sequence 620, App
11	20	100.0	86	9 US-09-392-598-274	Sequence 274, App
12	20	100.0	86	9 US-09-989-293A-274	Sequence 274, App
13	20	100.0	86	9 US-09-989-735-274	Sequence 274, App
14	20	100.0	86	9 US-09-990-444-274	Sequence 274, App
15	20	100.0	86	9 US-09-989-730-274	Sequence 274, App
16	20	100.0	86	9 US-09-990-436-274	Sequence 274, App
17	20	100.0	86	9 US-09-991-181-274	Sequence 274, App
18	20	100.0	86	9 US-09-993-687-274	Sequence 274, App
19	20	100.0	86	9 US-09-989-734-274	Sequence 274, App

ALIGNMENTS

RESULT 1

US-10-067-813-5
; Sequence 5, Application US/10067813
; Patent No. US20020156013A1
; GENERAL INFORMATION:
; APPLICANT: Renauld, Jean-Christophe
; APPLICANT: Louahed, Jamila
; APPLICANT: Grasso, Luigi
; APPLICANT: Levitt, Roy
; APPLICANT: Nicolaides, Nicholas
; TITLE OF INVENTION: Asthma Associated Factors as Targets for Treating
; TITLE OF INVENTION: Atopic Allergies and Related Disorders
; FILE REFERENCE: 036870-5071
; CURRENT APPLICATION NUMBER: US/10/067,813
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/157,247
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-813-5

Query Match 100.0%; Score 20; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 DRAT 4

10 DRAT 13

Db

10 DRAT 13

RESULT 2

US-10-043-487-532
; Sequence 532, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polyptei

;; TITLE OF INVENTION: mammalian polypeptides
;; FILE REFERENCE: B4778A
;; CURRENT APPLICATION NUMBER: US/10/043,487
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/261,130
;; PRIOR FILING DATE: 2001-01-12
;; NUMBER OF SEQ ID NOS: 561
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 532
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Shigella Flexneri
US-10-043-487-532

Query Match 100.0%; Score 20; DB 9; Length 26;

Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 11 DRAT 14

RESULT 3

US-10-142-231-37
;; Sequence 37, Application US/10142231
;; Publication No. US2003007796A1
;; GENERAL INFORMATION:

;; APPLICANT: Croteau, Rodney et al.

;; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES

;; FILE REFERENCE: 62773

;; CURRENT APPLICATION NUMBER: US/10/142,231

;; CURRENT FILING DATE: 2002-05-08

;; PRIOR APPLICATION NUMBER: 60/165,250

;; PRIOR FILING DATE: 1999-11-12

;; NUMBER OF SEQ ID NOS: 95

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 37

;; LENGTH: 50

;; TYPE: PRT

;; ORGANISM: Taxus cuspidata

US-10-142-231-37

Query Match

Best Local Similarity 100.0%; Score 20; DB 9; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 26 DRAT 29

RESULT 4

US-09-864-761-37634
;; Sequence 37634, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:

;; APPLICANT: Penn, Sharron G.

;; APPLICANT: Rank, David R.

;; APPLICANT: Hanzel, David K.

;; APPLICANT: Chen, Wensheng

;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

;; FILE REFERENCE: Acomica-X-1

;; CURRENT APPLICATION NUMBER: US/09/864,761

;; CURRENT FILING DATE: 2001-05-23

;; PRIOR APPLICATION NUMBER: US 60/180,312

;; PRIOR FILING DATE: 2000-02-04

;; PRIOR APPLICATION NUMBER: US 60/207,456

;; PRIOR FILING DATE: 2000-05-26

;; PRIOR APPLICATION NUMBER: US 09/632,366

;; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: GB 24263.6

;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 37634
;; LENGTH: 57
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL132772.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.7
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3
;; OTHER INFORMATION: SWISSPROT HIT: P41235, EVALUE 3.00e-30
;; OTHER INFORMATION: EST_HUMAN HIT: AU130255.1, EVALUE 6.00e-13
US-09-864-761-37634

Query Match 100.0%; Score 20; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 26 DRAT 29

RESULT 5

US-09-864-761-36163

;; Sequence 36163, Application US/09864761

;; Patent No. US20020048763A1

;; GENERAL INFORMATION:

;; APPLICANT: Penn, Sharron G.

;; APPLICANT: Rank, David R.

;; APPLICANT: Hanzel, David K.

;; APPLICANT: Chen, Wensheng

;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

;; FILE REFERENCE: Acomica-X-1

;; CURRENT APPLICATION NUMBER: US/09/864,761

;; CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36163
LENGTH: 66
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC012067.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.5
OTHER INFORMATION: SWISSPROT HIT: P15293, EVALUATE 5.00e-01
OTHER INFORMATION: EST_HUMAN HIT: BF364203.1, EVALUATE 9.00e-20
US-09-864-761-36163

Query Match 100.0%; Score 20; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 48 DRAT 51

RESULT 6

US-10-102-806-807
Sequence 807, Application US/10102806
Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103PIC1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 807
LENGTH: 72
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (7)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (18)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (21)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (42)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (54)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (61)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (65)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (67)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (68)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-807

Query Match 100.0%; Score 20; DB 9; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 29 DRAT 32

RESULT 7

US-09-864-761-34962
Sequence 34962, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312

;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 34962
;; LENGTH: 72
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL121755.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;; OTHER INFORMATION: EST HUMAN HIT: AW997712.1, EVALUATE 2.00e-25
;; OTHER INFORMATION: SWISSPROT HIT: P08547, EVALUATE 6.00e-21
US-09-864-761-34962

Query Match 100.0%; Score 20; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 52 DRAT 55

RESULT 8

US-09-864-761-34713
; Sequence 34713, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 34713
;; LENGTH: 78
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL121757.2
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 9.9
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.7
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;; OTHER INFORMATION: EST HUMAN HIT: AW997711.1, EVALUATE 7.00e-28
;; OTHER INFORMATION: SWISSPROT HIT: P08547, EVALUATE 2.00e-22
US-09-864-761-34713

Query Match 100.0%; Score 20; DB 10; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
DB 58 DRAT 61

RESULT 9

US-10-092-154-620
 ; Sequence 620, Application US/10092154
 ; Publication No. US20030054375A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC009C1
 ; CURRENT APPLICATION NUMBER: US/10/092,154
 ; CURRENT FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 2003
 ; Prior Application removed - See File Wrapper or Palm
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 620
 ; LENGTH: 84
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (37)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: misc_feature
 ; LOCATION: (83)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-10-092-154-620

Query Match 100.0%; Score 20; DB 9; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Length 84;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

||||
 1 DRAT 4

RESULT 10

US-09-764-847-620
 ; Sequence 620, Application US/09764847
 ; Patent No. US20020132767A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC009
 ; CURRENT APPLICATION NUMBER: US/09/764,847
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2003
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 620
 ; LENGTH: 84
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (37)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (83)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-764-847-620

Query Match 100.0%; Score 20; DB 10; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Length 84;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

||||
 1 DRAT 4

RESULT 11

US-09-992-598-274

; Sequence 274, Application US/0992598
 ; Patent No. US20020160384A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kijavini, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: ROY, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2730PIC20
 ; CURRENT APPLICATION NUMBER: US/09/992,598
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033

;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23

;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
|||
Db 74 DRAT 77

RESULT 12
US-09-989-293A-274
; Sequence 274, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444

; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.8e-02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 74 DRAT 77

RESULT 13

US-09-989-735-274

Sequence 274, Application US/09989735

Publication No. US20020193299A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnovers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730PIC61
 ; CURRENT APPLICATION NUMBER: US/09/989,735
 ; CURRENT FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734

;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 74 DRAT 77

RESULT 14
US-09-990-444-274
; Sequence 274, Application US/09990444
; Publication No. US20020193300A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C19
CURRENT APPLICATION NUMBER: US/09/990,444
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676

PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 1 DRAT 4

Db 74 DRAT 77

RESULT 15

US-09-989-730-274
Sequence 274, Application US/09989730
Publication No. US20020197674A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C69
CURRENT APPLICATION NUMBER: US/09/989,730
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861

; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 74 DRAT 77

Search completed: May 6, 2003, 15:04:53
Job time : 9.33333 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:30 ; Search time 6.4 Seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-851-422B-2

Perfect score: 20
Sequence: 1 DRAT 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	1	US-08-222-851-18
2	20	100.0	16	2	US-08-797-842-3
3	20	100.0	19	4	US-08-952-127-16
4	20	100.0	21	4	US-08-952-127-13
5	20	100.0	29	4	US-08-817-811-83
6	20	100.0	40	2	US-08-797-842-11
7	20	100.0	40	2	US-08-797-842-12
8	20	100.0	44	1	US-07-956-700B-60
9	20	100.0	44	1	US-08-476-537-60
10	20	100.0	44	1	US-08-485-607-60
11	20	100.0	44	2	US-08-475-879-60
12	20	100.0	44	4	US-09-433-043B-60
13	20	100.0	70	1	US-07-737-736B-3
14	20	100.0	89	4	US-09-187-789-15
15	20	100.0	89	4	US-09-139-600-10
16	20	100.0	102	2	US-08-480-473B-51
17	20	100.0	102	3	US-08-915-213-51
18	20	100.0	102	4	US-09-235-217-51
19	20	100.0	105	1	US-08-276-852-93
20	20	100.0	105	1	US-08-899-575-93
21	20	100.0	105	1	US-08-899-575-93
22	20	100.0	105	5	PCT-US95-08743-93
23	20	100.0	107	1	US-08-276-852-114
24	20	100.0	107	1	US-08-899-575-114
25	20	100.0	107	1	US-08-899-575-114
26	20	100.0	107	5	PCT-US95-08743-114
27	20	100.0	108	1	US-08-276-852-98

28 20 100.0 108 1 US-08-899-575-98 Sequence 98, Appl
29 20 100.0 108 1 US-08-899-575-98 Sequence 98, Appl
30 20 100.0 108 5 PCT-US95-08743-98 Sequence 98, Appl
31 20 100.0 109 4 US-09-419-459-5 Sequence 5, Appl
32 20 100.0 109 4 US-09-419-459-6 Sequence 6, Appl
33 20 100.0 110 4 US-09-187-859-22 Sequence 22, Appl
34 20 100.0 125 3 US-08-959-382-4 Sequence 4, Appl
35 20 100.0 132 4 US-09-215-221-28 Sequence 28, Appl
36 20 100.0 133 4 US-08-994-962-3 Sequence 3, Appl
37 20 100.0 155 2 US-08-576-626A-49 Sequence 49, Appl
38 20 100.0 163 1 US-07-991-867B-7 Sequence 7, Appl
39 20 100.0 163 1 US-08-107-755A-7 Sequence 7, Appl
40 20 100.0 163 2 US-08-544-332-7 Sequence 7, Appl
41 20 100.0 163 4 US-09-370-861A-7 Sequence 7, Appl
42 20 100.0 184 4 US-09-384-162-12 Sequence 12, Appl
43 20 100.0 215 4 US-08-861-774E-36 Sequence 36, Appl
44 20 100.0 222 6 5223425-6 Patent No. 5223425
45 20 100.0 226 2 US-08-658-277-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-222-851-18
; Sequence 18, Application US/08222851
; Patent No. 5723128
; GENERAL INFORMATION:
; APPLICANT: CLAYBERGER, CAROL A.
; APPLICANT: KRENSKY, ALAN M.
; APPLICANT: PARHAM, PETER
; TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")
; TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,851
FILING DATE: 05-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 28600-20200.22
TELEPHONE: (202) 887-1500
TELEFAX: (202) 494-0792
TELEX: 90-4030 MRSNFOERSMWH
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-851-18

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4

Db 1 DRAT 4

```
RESULT 2
US-08-797-842-3
; Sequence 3, Application US/08797842
; Patent No. 5932706
; GENERAL INFORMATION:
; APPLICANT: Mertens, Koenraad et al
; TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
; TITLE OF INVENTION: their use for isolating intact protein, haemostatic composition
; TITLE OF INVENTION: of proteolytic cleavage products of the protein
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michaelson and Wallace
; STREET: Parkway 109 Office Center, 328 Newman Springs
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11
; SOFTWARE: Microsoft Word for Windows 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,842
; FILING DATE: 10-Feb-1997
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 530
; FILING DATE: 10-Feb-1997
; APPLICATION NUMBER: 08/381,891
; FILING DATE: February 8, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Stitching-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 530-6571
; TELEFAX: (908) 530-6584
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-797-842-3
Query Match 100.0%; Score 20; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 13 DRAT 16

RESULT 3
US-08-952-127-16
; Sequence 16, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Tagle, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,127
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
```

```
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,127
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-952-127-16
Query Match 100.0%; Score 20; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 7 DRAT 10

RESULT 4
US-08-952-127-13
; Sequence 13, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Tagle, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,127
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-952-127-13

Query Match 100.0%; Score 20; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||
Db 7 DRAT 10

RESULT 5

US-08-817-811-83
Sequence 83, Application US/08817811
Patent No. 6174528

GENERAL INFORMATION:

APPLICANT: Cooper, Juan A.
APPLICANT: Relf, Wendy A.
APPLICANT: Good, Michael F.
APPLICANT: Saul, Allan J.
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
TITLE OF INVENTION: COMPRISING SAME
NUMBER OF SEQUENCES: 97
CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,811
FILING DATE: 14-APR-1997
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 96/11944
FILING DATE: 25-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: FBRC:005
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-817-811-83

Query Match 100.0%; Score 20; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||
Db 8 DRAT 11

RESULT 6

US-08-797-842-11
Sequence 11, Application US/08797842
Patent No. 5932706

GENERAL INFORMATION:

APPLICANT: Mertens, Koenraad et al
TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
their use for isolating intact protein, haemostatic composition,
TITLE OF INVENTION: of proteolytic cleavage products of the protein
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michaelson and Wallace
STREET: Parkway 109 Office Center, 328 Newman Springs
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11
SOFTWARE: Microsoft Word for Windows 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,842
FILING DATE: 10-Feb-1997
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/381,891
FILING DATE: February 8, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Stitching-5
TELEPHONE: (908)530-6671
TELEFAX: (908)530-6584
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-797-842-11

Query Match 100.0%; Score 20; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||
Db 25 DRAT 28

RESULT 7

US-08-797-842-12
Sequence 12, Application US/08797842
Patent No. 5932706

GENERAL INFORMATION:

APPLICANT: Mertens, Koenraad et al
TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
their use for isolating intact protein, haemostatic composition,
TITLE OF INVENTION: of proteolytic cleavage products of the protein
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Michaelson and Wallace
STREET: Parkway 109 Office Center, 328 Newman Springs
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette
COMPUTER: IBM PC

OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11
SOFTWARE: Microsoft Word for Windows 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,842
FILING DATE: 10-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/381,891
FILING DATE: February 8, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Stitching-S
TELEPHONE: (908)530-6671
TELEFAX: (908)530-6584
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-797-842-12

Query Match 100.0%; Score 20; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||
Db 14 DRAT 17

RESULT 8
US-07-956-700B-60
Sequence 60, Application US/07956700B
Patent No. 5535092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5539092th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5539092thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear

MOLECULE TYPE: Peptide
US-07-956-700B-60
Query Match 100.0%; Score 20; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRAT 4
|||
Db 25 DRAT 28
RESULT 9
US-08-476-537-60
Sequence 60, Application US/08476537
Patent No. 5756290
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5756290th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,537
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5756290thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-476-537-60

Query Match 100.0%; Score 20; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
|||
Db 25 DRAT 28

RESULT 10
US-08-485-607-60
Sequence 60, Application US/08485607
Patent No. 5792627
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-485-607-60

Query Match 100.0%; Score 20; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 25 DRAT 28

RESULT 11
US-08-475-879-60
Sequence 60, Application US/08475879
Patent No. 5972644
Patent No. 5972644 5786170
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5972644 5786170th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER: 07/956,700

FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-475-879-60

Query Match 100.0%; Score 20; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 25 DRAT 28

RESULT 12
US-09-433-043B-60
Sequence 60, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 60
LENGTH: 44
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-433-043B-60

Query Match 100.0%; Score 20; DB 4; Length 44;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 25 DRAT 28

RESULT 13
US-07-737-736B-3
Sequence 3, Application US/07737736B
Patent No. 5260199
GENERAL INFORMATION:
APPLICANT: Deluca, Hector F.
APPLICANT: Ross, Troy K.
APPLICANT: Prah, Jean M.
TITLE OF INVENTION: Method Of Producing
TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/737,736B
FILING DATE: 19910730
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Carl R.
REGISTRATION NUMBER: 29,437
REFERENCE/DOCKET NUMBER: 96-296-2185-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 414-277-5715
TELEFAX: 414-277-5774
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chicken
PUBLICATION INFORMATION:
AUTHORS: McDonnell, Donald P.
AUTHORS: Mangelsdorf, David J.
AUTHORS: Pike, J. W.
AUTHORS: Hausler, Mark R.
AUTHORS: O'Malley, Bert W.
TITLE: Molecular Cloning of Complementary DNA
TITLE: Encoding the Avian Receptor for Vitamin D
JOURNAL: Science
VOLUME: 235
PAGES: 1214-1217
DATE: March 6-1987
US-07-737-736B-3

Query Match 100.0%; Score 20; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 8 DRAT 11

RESULT 14
US-09-187-789-15
; Sequence 15, Application US/09187789
; Patent No. 6340740
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434C1
; CURRENT APPLICATION NUMBER: US/09/187,789
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 89
; TYPE: PRT

; ORGANISM: Mus musculus
US-09-187-789-15
Query Match 100.0%; Score 20; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 75 DRAT 78

RESULT 15
US-09-139-600-10
; Sequence 10, Application US/09139600
; Patent No. 6432628
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; TITLE OF INVENTION: AND METHOD OF USE
; FILE REFERENCE: 480140.434
; CURRENT APPLICATION NUMBER: US/09/139,600
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-139-600-10

Query Match 100.0%; Score 20; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRAT 4
Db 75 DRAT 78

Search completed: May 6, 2003, 15:04:06
Job time : 7.4 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:44 ; Search time 18.2 Seconds
(without alignments)
29.286 Million cell updates/sec

Title: US-09-851-422B-2

Perfect score: 20

Sequence: 1 DRAT 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	16 AAR83079	Class I MHC antigen
2	20	100.0	12	14 AAR30472	ICAM-3 peptide fra
3	20	100.0	12	21 AAB40057	Anti-hfl12 antibod
4	20	100.0	16	15 AAR50113	Peptide reactive w
5	20	100.0	19	18 AAW36170	Mouse ATM gene pro
6	20	100.0	19	20 AAY01771	M-ras derived anti
7	20	100.0	21	18 AAW07660	Human ATM gene pro
8	20	100.0	21	22 AAB87221	Breast-cancer asso
9	20	100.0	28	17 AAR97470	Chimaeric peptide
10	20	100.0	28	17 AAR91391	MEH1 library deriv

11.	20	100.0	36	22 AAG73731	Human colon cancer
12	20	100.0	37	21 AAY93158	Human vitamin D re
13	20	100.0	37	21 AAY64710	Human 5' EST relat
14	20	100.0	38	21 AAG56794	Arabidopsis thalia
15	20	100.0	38	22 AAM87857	Human immune/haema
16	20	100.0	40	21 AAB32462	Human secreted pro
17	20	100.0	43	22 ABB11187	Human pheromone re
18	20	100.0	44	20 AAW74045	Human D2H binding
19	20	100.0	46	12 AAR10755	Non-A non-B hepati
20	20	100.0	50	21 AAG07777	Arabidopsis thalia
21	20	100.0	50	22 AAU47989	Propionibacterium
22	20	100.0	50	22 AAU02815	Taxus cuspidata RT
23	20	100.0	52	22 AAU52605	Propionibacterium
24	20	100.0	53	22 AAU56447	Propionibacterium
25	20	100.0	54	22 AAU55929	Propionibacterium
26	20	100.0	54	22 AAU60960	Propionibacterium
27	20	100.0	54	22 AAU62928	Propionibacterium
28	20	100.0	56	22 AAU54832	Propionibacterium
29	20	100.0	57	22 AAU67185	Propionibacterium
30	20	100.0	57	22 ABB31792	Peptide #4443 enco
31	20	100.0	57	22 ABB37021	Peptide #4527 enco
32	20	100.0	57	22 ABB16368	Human nervous syst
33	20	100.0	57	22 ABB22336	Protein #4335 enco
34	20	100.0	57	22 AAM57748	Human brain expres
35	20	100.0	57	22 AAM70162	Human bone marrow
36	20	100.0	57	22 AAM17990	Peptide #4424 enco
37	20	100.0	57	22 AAM30496	Peptide #4533 enco
38	20	100.0	57	22 AAM05627	Peptide #4309 enco
39	20	100.0	57	23 AAG39799	Human peptide enco
40	20	100.0	57	23 ABP42618	Human ovarian anti
41	20	100.0	58	22 AAG75801	Human colon cancer
42	20	100.0	59	21 AAB40924	Human ORF ORF688
43	20	100.0	59	21 AAG01585	Human secreted pro
44	20	100.0	59	23 ABP31081	Human peptidase-li
45	20	100.0	60	22 AAU67029	Propionibacterium

ALIGNMENTS

RESULT 1
AAR83079
ID AAR83079 standard; peptide; 4 AA.
AC AAR83079;
XX
XX
DT 16-MAY-1996 (first entry)
XX
DE Class I MHC antigen alpha-1 domain tetrapeptide #3.
XX
XX
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC.
XX
OS Synthetic.
XX
PN WO9526979-A1.
XX
PD 12-OCT-1995.
XX
PF 05-APR-1995; 95WO-US04349.
XX
PR 05-APR-1994; 94US-0222851.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Clayberger C, Krensky AM, Parham P;
XX
DR WPI; 1995-358582/46.
XX
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host

XX PS Disclosure; Page 10; 80pp; English.

XX CC AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments of class I major histocompatibility complex (MHC) antigens. These sequences can be used to extend the period of acceptance by a recipient of a transplant from an MHC unmatched donor. The peptides are administered to a patient in conjunction with a subtherapeutic amount of an immunosuppressant. This is administered to the patient for a limited period of time (compared to the lifetime administration for current treatments). The peptides particularly modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs) of the patient.

XX CC

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 16; Length 4;

Best Local Similarity 100.0%; Pred. NO. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 1 DRAT 4

RESULT 2

AAR30472

ID AAR30472 standard; peptide; 12 AA.

XX AC AAR30472;

XX DT 12-MAY-1993 (first entry)

XX DE ICAM-3 peptide fragment NK-10.

XX KW Intercellular adhesion molecule-3; ICAM-3; ICAM-1; ICAM-2; homology; NK-17; Ig domain; NK-10; domain 4; probe; leukocyte; lymphocyte; macrophage; neutrophil; inflammation; immune response.

XX OS Homo sapiens.

XX PN W09222323-A.

XX PD 23-DEC-1992.

XX PF 11-JUN-1992; 92WO-US04896.

XX PR 11-JUN-1991; 91US-0712879.

XX PA (BLOO-) CENT BLOOD RES INC.

XX PI DeFougerolles AR, Springer TA;

XX DR WPI; 1993-017908/02.

XX PT Inter-cellular adhesion molecule-3 inhibiting granulocyte, lymphocyte and macrophage adhesion - for treating inflammation, AIDS, asthma, auto-immune thyroiditis, multiple sclerosis, ARDS etc.

XX PS Disclosure; Page 75; 123pp; English.

XX CC The sequences given in AAR30472-73 represent fragments of intercellular adhesion molecule-3 (ICAM-3). Comparisons of these sequences with known sequences from ICAM-1 and ICAM-2 reveals a high degree of homology between NK-17 and sequences shown in the first Ig domain of ICAM-2. NK-10 peptide shows weak homology to sequences within domain 4 of ICAM-1. These peptide sequences were used to design probes for the isolation of the cDNA encoding ICAM-3 (see also AAQ33110-11). ICAM-3 is involved in the process by which populations of leukocytes recognise and adhere to cellular substrates. ICAM-3 mediates cellular interactions with other lymphocytes, macrophages and neutrophils at the site of inflammation and sites of immune responses.

XX CC

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 20; DB 14; Length 12;

Best Local Similarity 100.0%; Pred. NO. 59;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4

Db 3 DRAT 6

RESULT 3

AAB40057

ID AAB40057 standard; Peptide; 12 AA.

XX AC AAB40057;

XX DT 05-FEB-2001 (first entry)

XX DE Anti-HIL12 antibody light chain CDR3 amino acid sequence SEQ ID 573.

XX KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory; complementarity determining region; CDR; antirheumatic; antiarthritic; antisclerotic; neuroprotective; antiposrotic; antiasthmatic; cardiant; antiparasitic; antibacterial; immunosuppressive; Crohn's disease; multiple sclerosis; rheumatoid arthritis.

XX OS Homo sapiens.

XX PN W0200056772-A1.

XX PD 28-SEP-2000.

XX PF 24-MAR-2000; 2000WO-US07946.

XX PR 25-MAR-1999; 99US-0126603.

XX PA (BADI) BASF AG.

XX PA (GEMY) GENETICS INST INC.

XX PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M; Kalmakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A; Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR; Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;

XX DR WPI; 2000-638250/61.

XX PT New human antibody specific for human interleukin-12 (IL-12) used to treat disorders characterized by aberrant IL-12 expression e.g. Crohn's disease and multiple sclerosis -

XX PS Claim 32; Figure 2H; 377pp; English.

XX CC This invention relates to a new human antibody specific for human interleukin-12 (IL-12). The invention also includes antigen binding portions that bind to IL-12. Sequences AAB39485-B39516 represent human anti-IL-12 antibody heavy and light chain complementarity determining region (CDR) amino acid sequences, and also includes variable region amino acid sequences. Other variable region amino acid sequences are given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771 represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063 represent other CDR sequences. Light chain CDR3 consensus sequences are given in AAB40064-B40067. Primers used in the identification and construction of the antibodies of the invention are given in AAC61062-C61071. The antibody of the invention is a neutralising antibody and has antirheumatic; antiarthritic; antisclerotic; antiinflammatory; neuroprotective; antiposrotic; antiasthmatic; cardiant; antiparasitic; antibacterial and immunosuppressive activity. The antibodies or antigen-binding fragments are useful in the treatment of disorders associated with detrimental release of human IL-12, especially Crohn's disease, multiple sclerosis and rheumatoid arthritis. They can also be used in the manufacture of a pharmaceutical composition to treat human IL-12 disorders.

XX CC

```

XX SQ Sequence 12 AA;
Query Match 100.0%; Score 20; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRAT 4
DB 4 DRAT 7

RESULT 4
AAR50113
ID AAR50113 standard; peptide; 16 AA.
XX AC AAR50113;
XX 07-OCT-1994 (first entry)
XX Peptide reactive with antibody specific for factor IX.
XX Polypeptide; peptide; oligopeptide; factor IX; haemostatic protein;
KW antibody; treatment; haemostatic disease; thromboembolic disease.
XX Homo sapiens.
XX WO9405692-A.
XX 17-MAR-1994.
XX 26-AUG-1993; 93WO-NL00174.
XX 27-AUG-1992; 92EP-0202615.
XX (BLOE-) STICHTING CENT LAB VAN DE BLOEDTRANSFUSI.
XX Mertens K, Van Mourik JA;
XX WPI; 1994-101117/12.
XX Isolating haemostatic proteins free of proteolytic degradation
PT products - using new antibodies which distinguish between intact
PT and cleaved forms of the protein, useful in treatment of
PT haemostatic diseases
XX Claim 6; Page 25; 32pp; English.
XX The polypeptide can be used to screen antibodies to select
CC antibodies which are specific for a haemostatic protein and can
CC substantially distinguish between intact and cleaved species of the
CC haemostatic protein. The polypeptide corresponds to amino acid
CC residues 320-335 of factor IX. Haemostatic proteins are used in the
CC treatment of haemostatic e.g. thromboembolic diseases. See
CC AAR50111-R50114.
XX SQ Sequence 16 AA;
Query Match 100.0%; Score 20; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRAT 4
DB 13 DRAT 16

RESULT 5
AAW36170
ID AAW36170 standard; Peptide; 19 AA.
XX AC AAW36170;
XX

30-MAR-1998 (first entry)
Mouse ATM gene product epitope 1.
Ataxia-telangiectasia; A-T; mutated; ATM; 11q22-23; signal transduction;
DNA damage; cell cycle control; screening; gene therapy; catalytic;
phosphatidylinositol-3 kinase; PI-3; cancer predisposition; carrier;
untranslated region; UTR.
Mus sp.
WO9636695-A1.
21-NOV-1996.
16-MAY-1996; 96WO-US07040.
28-JUL-1995; 95US-0508836.
16-MAY-1995; 95US-0441822.
21-JUN-1995; 95US-0493092.
(UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Collins FS, Shiloh Y, Tagle DA;
WPI; 1997-012074/01.
New gene ATM implicated in ataxia-telangiectasia and related protein
carriers
Claim 24; Page 105; 127pp; English.
AAW36170-72 are mouse ATM protein epitopes used to generate polyclonal
antibodies. Antibodies raised against the ATM protein detected
mono-specifically a high molecular weight of the expected size of 350
kDa on Western blots of protein lysates derived from fibroblast and
lympho- blastoid cell lines. Because of the high frequency of truncation
mutations in the ATM gene, mutated ATM protein can be identified if such
proteins are stable. Mutations in the ATM gene cause ataxia-
telangiectasia (A-T), a progressive genetic disorder affecting the
central nervous and immune systems. The ATM gene, located at chromosome
11q22-23, is probably involved in a novel signal transduction system
that links DNA damage surveillance to cell cycle control. The ATM gene
product (AAW07655) has a highly conserved C-terminal region showing high
sequence homology to the catalytic domain of phosphatidylinositol-3
kinases. A-T mutations affect a variety of tissues and lead to cancer
predisposition. Identification of A-T carriers, by analysis at nucleic
acid or protein levels, allows better supervision and treatment of such
subjects who are at increased risk of developing cancer and are
particularly sensitive to radiation. The transgenic animals and
transformed cells are useful as models of the human disease. Also viral
vectors expressing the ATM protein can be used in gene therapy of A-T.
XX SQ Sequence 19 AA;
Query Match 100.0%; Score 20; DB 18; Length 19;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRAT 4
DB 7 DRAT 10

RESULT 6
AAW01771
ID AAW01771 standard; Peptide; 19 AA.
XX AC AAW01771;
XX 28-JUN-1999 (first entry)

```

XX M-ras derived antigenic peptide used to raise antibodies.
DE
XX
XX M-ras; interleukin-9 signaling pathway; pathogenesis; atopic allergy;
KW asthma; leukemia; lymphoma; tumor; T-cell lymphoma; Hodgkin's disease;
KW mycosis fungoides.
XX
XX Unidentified.
OS
XX WO9914242-A2.
PN
XX
XX 25-MAR-1999.
PD
XX
XX 18-SEP-1998; 98WO-US19626.
PF
XX 19-SEP-1997; 97US-0059509.
PR
XX (MAGA-) MAGAININ PHARM INC.
PA
XX Grasso L, Levitt R, Louahed J, Nicolaides N, Renauld J;
PI WPI; 1999-244027/20.
XX Nucleic acid encoding human or murine M-Ras
DR
XX Disclosure; Page 17; 69pp; English.
XX
XX The present sequence represents a peptide derived from M-Ras
CC protein, which was used to raise antibodies. M-Ras, a part of
CC the interleukin-9 signaling pathway, is involved in pathogenesis of
CC atopic allergy, asthma and similar conditions, some leukemias and
CC lymphomas and tumors. Agents that downregulate M-Ras are used to
CC treat these conditions, particularly T-cell lymphoma or leukemia;
CC Hodgkin's disease and mycosis fungoides. Measuring levels of M-Ras
CC is used to diagnose, or monitor, these conditions (an elevated level
CC being indicative of disease).
XX
XX Sequence 19 AA;
SQ
Query Match 100.0%; Score 20; DB 20; Length 19;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRAT 4
DB 10 DRAT 13
RESULT 7
AAW07660
ID AAW07660 standard; Peptide; 21 AA.
XX
XX AAW07660;
AC
XX 30-MAR-1998 (first entry)
DT
XX Human ATM gene product epitope 5.
DE
XX Ataxia-telangiectasia; A-T; mutated; ATM; 11q22-23; signal transduction;
KW DNA damage; cell cycle control; screening; gene therapy; catalytic;
KW phosphatidylinositol-3 kinase; PI-3; cancer predisposition; carrier;
KW untranslated region; UTR.
XX
XX Homo sapiens.
OS
XX WO9636695-A1.
PN
XX 21-NOV-1996.
PD
XX 16-MAY-1996; 96WO-US07040.
XX
XX 28-JUL-1995; 95US-0508836.
PR 16-MAY-1995; 95US-0441822.
PR

PR 21-JUN-1995; 95US-0493092.
XX (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Collins FS, Shiloh Y, Tagle DA;
PI WPI; 1997-012074/01.
XX
XX New gene ATM implicated in ataxia-telangiectasia and related protein
PT carriers
PT - useful in screening methods, partic. for identifying disease
XX
XX Example 5; Page 104; 127pp; English.
XX
XX AAW07656-62 are human ATM protein epitopes used to generate polyclonal
CC and monoclonal antibodies. Antibodies raised against the ATM protein
CC detected mono-specifically a high molecular weight of the expected size
CC of 350 kDa on Western blots of protein lysates derived from fibroblast
CC and lympho- blastoid cell lines. Because of the high frequency of
CC truncation mutations in the ATM gene, mutated ATM protein can be
CC identified if such proteins are stable. Mutations in the ATM gene cause
CC ataxia- telangiectasia (A-T), a progressive genetic disorder affecting
CC the central nervous and immune systems. The ATM gene, located at
CC chromosome 11q22-23, is probably involved in a novel signal transduction
CC system that links DNA damage surveillance to cell cycle control. The ATM
CC gene product (AAW07655) has a highly conserved C-terminal region showing
CC high sequence homology to the catalytic domain of phosphatidylinositol-3
CC kinases. A-T mutations affect a variety of tissues and lead to cancer
CC predisposition. Identification of A-T carriers, by analysis at nucleic
CC acid or protein levels, allows better supervision and treatment of such
CC subjects who are at increased risk of developing cancer and are
CC particularly sensitive to radiation. The transgenic animals and
CC transformed cells are useful as models of the human disease. Also viral
CC vectors expressing the ATM protein can be used in gene therapy of A-T.
XX
XX Sequence 21 AA;
SQ
Query Match 100.0%; Score 20; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRAT 4
DB 7 DRAT 10
RESULT 8
AAB87221
ID AAB87221 standard; Peptide; 21 AA.
XX
XX AAB87221;
AC
XX 10-MAY-2001 (first entry)
DT
XX Breast-cancer associated protein isoform BPI-20 peptide #2.
DE
XX Human; breast cancer; breast cancer associated protein isoform; BPI;
KW breast cancer associated feature; BF; diagnosis; cytostatic.
XX
XX Homo sapiens.
OS
XX WO200113117-A2.
PN
XX 22-FEB-2001.
PD
XX 14-AUG-2000; 2000WO-GB03143.
XX
XX 13-AUG-1999; 99GB-0019258.
PR 30-MAR-2000; 2000GB-0007754.
XX
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX

DR WPI; 1996-151391/15.
 XX Synthetic random nucleotide sequences encoding ligand binding
 PT domains - identified by screening library of vectors or peptide(s)
 PT and useful for gene therapy and diagnosis
 XX Claim 70; Page 196; 224pp; English.
 XX The sequences given in AAR91378-414 represent peptides that bind to
 CC highly specific DNA binding domains (HSDB)'s. These sequences were
 CC tested for binding to the H2-kappa-B oligonucleotide which contains
 CC the NF-kappa-B binding site, and comprises the sequence given in
 CC AAT13579. These peptides may be used in a composition for diagnosis and
 CC gene therapy and for modifying the transcription or activity of a gene.
 XX
 SQ Sequence 28 AA;
 Query Match 100.0%; Score 20; DB 17; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRAT 4
 DB 12 DRAT 15
 RESULT 11
 AAG73731
 ID AAG73731 standard; Protein; 36 AA.
 AC AAG73731;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:4495.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 XX WO200122920-A2.
 XX
 XX 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000NO-US26524.
 XX
 XX 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-235357/24.
 DR N-PSDB; AAH33162.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PT
 PS Claim 11; Page 6295; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,

CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 36 AA;
 Query Match 100.0%; Score 20; DB 22; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRAT 4
 DB 28 DRAT 31
 RESULT 12
 AAY93158
 ID AAY93158 standard; peptide; 37 AA.
 XX
 AC AAY93158;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE Human vitamin D receptor amino acids 18-54.
 XX
 KW Cytostatic; mutant; mutin; oestrogen receptor-alpha; zinc finger;
 KW oestrogen receptor-beta; vitamin D receptor; retinoic acid receptor;
 KW thyroid hormone receptor; proliferation; motility; invasiveness; cancer;
 KW breast; ovary.
 XX
 OS Homo sapiens.
 XX
 PN WO200035955-A1.
 XX
 XX 22-JUN-2000.
 XX
 XX 17-DEC-1999; 99WO-FR03173.
 PF
 XX 17-DEC-1998; 98FR-0015922.
 PR
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA
 XX Garcia M, Platet N, Rochefort H;
 XX WPI; 2000-442369/38.
 DR
 XX Composition containing nucleic acid encoding variant estrogen receptor
 PT reduces proliferation, motility and invasiveness of cancer cells,
 PT useful for treating or preventing cancer -
 XX
 XX Example 2; Fig 3; 33pp; French.
 XX
 CC The invention relates to a pharmaceutical composition containing at
 CC least a nucleic acid encoding a variant estrogen receptor that
 CC comprises part of the receptor's C region (containing the first zinc
 CC finger) but which lacks the ability to bind estrogen and/or AP-2 (a
 CC transcriptional activator). The variation is especially in the region
 CC comprising amino acid 179-215 of the alpha estrogen receptor. Similar
 CC variations may be generated in the homologous of other estrogen-type
 CC receptor such as oestrogen receptor-beta, vitamin D receptor, retinoic
 CC acid receptor or thyroid hormone receptor (see peptides
 CC AA93155-Y93159). The receptor variants reduce proliferation, motility
 CC and invasiveness of cancer cells. They can therefore be used to treat or
 CC prevent cancer, specifically of the breast or ovary.
 XX
 SQ Sequence 37 AA;
 Query Match 100.0%; Score 20; DB 21; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 12 DRAT 15

RESULT 13
AA64710
ID AAY64710 standard; Protein; 37 AA.

AC AAY64710;

XX 01-FEB-2000 (first entry)

XX Human 5' EST related polypeptide SEQ ID NO:871.

XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.

XX Homo sapiens.

XX WO953051-A2.

XX 21-OCT-1999.

XX 09-APR-1999; 99WO-IB00712.

XX 09-APR-1998; 98US-0057719.

XX 28-APR-1998; 98US-0069047.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-038446/03.

XX N-PSDB; AA242324.

XX Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures

XX Claim 3; Page 618; 837pp; English.

XX AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AA64651 to
CC AAY65438 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.

XX SQ Sequence 37 AA;

Query Match 100.0%; Score 20; DB 21; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
Db 33 DRAT 36

RESULT 14

AA656794

ID AAG56794 standard; Protein; 38 AA.

XX

XX AAG56794;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 73090.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 16-JUN-1999; 99US-0139453.

XX 17-JUN-1999; 99US-0139492.

ID AM87857 standard; Protein; 38 AA.
XX
AC AM87857;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:15450.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX WO200157182-A2.
PN
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-0231143.
PR 08-SEP-2000; 2000US-0231143.
PR 08-SEP-2000; 2000US-0231144.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250191.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251858.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) *HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 DR N-PSDB; AAK60638.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 11; SEQ ID NO 15450; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity; and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 38 AA;

Query Match 100.0%; Score 20; DB 22; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAT 4
 |||
 Db 6 DRAT 9

Search completed: May 6, 2003, 15:01:02
 Job time : 19.2 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:45 ; Search time 24.85 Seconds
(without alignments)
58.041 Million cell updates/sec

Title: US-09-851-422B-8

Perfect score: 39
Sequence: 1 VPHNESE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_prodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	361	11 Q9DCH4	Q9dch4 mus musculus
2	36	92.3	315	10 Q9LE16	Q9le16 arabidopsis
3	36	92.3	315	10 Q93254	Q93254 arabidopsis
4	36	92.3	320	10 Q9C928	Q9c928 arabidopsis
5	35	89.7	1344	12 Q69071	Q69071 human herpe
6	34	87.2	293	10 Q84H4	Q84h4 arabidopsis
7	34	87.2	293	10 Q04202	Q04202 arabidopsis
8	34	87.2	302	3 Q43660	Q43660 schizosacch
9	34	87.2	377	11 Q921F7	Q921f7 mus musculus
10	33	84.6	381	16 Q8U8U0	Q8u8u0 agrobacteri
11	33	84.6	510	5 Q97251	Q97251 plasmodium
12	32	82.1	115	5 Q9V53	Q9v53 drosophila
13	32	82.1	332	10 Q9LES7	Q9les7 arabidopsis
14	32	82.1	369	13 Q9J345	Q9j345 gallus gall
15	32	82.1	372	10 Q05699	Q05699 nicotiana t
16	32	82.1	1639	5 Q9VNF2	Q9vnf2 drosophila

17	32	82.1	1689	12 Q9IWI4	Q9iwi4 crimean-con
18	32	82.1	1689	12 Q8ZG0	Q8zgo crimean-con
19	32	82.1	1819	16 Q9THX3	Q9thx3 clostridium
20	32	82.1	3021	12 Q81258	Q81258 hepatitis c
21	31	79.5	249	10 Q9C557	Q9c557 arabidopsis
22	31	79.5	253	5 Q8T4L9	Q8t4l9 drosophila
23	31	79.5	272	5 Q9XW43	Q9xw43 caenorhabdi
24	31	79.5	316	16 Q8ZAG5	Q8zag5 yersinia pe
25	31	79.5	316	16 Q9LSP6	Q9lsp6 salmonella
26	31	79.5	374	5 Q19006	Q19006 caenorhabdi
27	31	79.5	382	5 P91040	P91040 caenorhabdi
28	31	79.5	457	5 Q9GSL2	Q9gsl2 ciona intes
29	31	79.5	534	5 Q9VSD7	Q9vsd7 drosophila
30	31	79.5	548	5 Q9XUC0	Q9xuc0 caenorhabdi
31	31	79.5	607	16 Q98AV5	Q98av5 rhizobium 1
32	31	79.5	648	2 Q93FT6	Q93ft6 cowdria rum
33	31	79.5	792	17 Q97BB5	Q97bb5 thermoplasma
34	31	79.5	1068	10 Q9MAQ7	Q9maq7 arabidopsis
35	31	79.5	1098	13 Q91951	Q91951 coturnix co
36	31	79.5	1274	2 Q45563	Q45563 bacillus su
37	31	79.5	1391	3 Q8XOV7	Q8xov7 neurospora
38	31	79.5	1467	10 Q9M033	Q9m033 arabidopsis
39	31	79.5	1474	5 Q8T4M0	Q8t4m0 drosophila
40	31	79.5	1503	5 Q8T4L8	Q8t4l8 drosophila
41	31	79.5	1509	5 Q95P10	Q95p10 drosophila
42	30	76.9	84	11 Q9CUV6	Q9cuv6 mus musculu
43	30	76.9	97	2 Q05562	Q05562 mycobacteri
44	30	76.9	143	2 Q9Z5H0	Q9z5h0 mycobacteri
45	30	76.9	193	10 Q24284	Q24284 phoenix rec

ALIGNMENTS

RESULT 1

Q9DCH4	Q9DCH4	PRELIMINARY;	PRT;	361 AA.
ID	Q9DCH4;			
AC	Q9DCH4;			
DC	01-JUN-2001 (TREMREL. 17, Created)			
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)			
DT	01-JUN-2002 (TREMREL. 21, Last annotation update)			
DE	0610037M02rik protein.			
GN	E1F3S5 OR 0610037M02RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;			
RC	MEDLINE=21085560; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,			
RA	Hayashizaki Y.;			
RL	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK002778; BAB22352.1; -.			
DR	MGD; MGI:1913335; E1f3s5.			

```

DR InterPro: IPR000555; Mov34.
DR InterPro: IPR003640; Mov34.2.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF01398; Mov34.1.
DR PRINTS: PR01217; PRICHEXTENS.
DR PRODOM: PD005425; Mov34_2; 1.
DR SMART: SM00232; JAB_MFN; 1.
SQ SEQUENCE 361 AA; 38000 MW; 8AA985DCD248E3DE CRC64;

Query Match 100.0%; Score 39; DB 11; Length 361;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
DB 141 VPHNESE 147
|||||

RESULT 2
Q9LE16 PRELIMINARY; PRT; 315 AA.
AC Q9LE16;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Similarity to unknown protein (WJL13.12 protein).
GN WJL13.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., C.W.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III P1 WJL13 genomic sequence.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB022218; BAB02375.1; --
DR EMBL: AAC24081; AAP35412.1; --
SQ SEQUENCE 315 AA; 36350 MW; 2D4327073B9883F2 CRC64;

Query Match 92.3%; Score 36; DB 10; Length 315;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
DB 218 VPHNESE 224
|||||

RESULT 3
Q93Z54 PRELIMINARY; PRT; 315 AA.
AC Q93Z54;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AT3G15460/WJL13.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY058120; AAL25537.1; --
SQ SEQUENCE 315 AA; 36336 MW; 5B16B594E985BF08 CRC64;

Query Match 92.3%; Score 36; DB 10; Length 315;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
DB 218 VPHNESE 224
|||||

RESULT 4
Q9C928 PRELIMINARY; PRT; 320 AA.
AC Q9C928;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 36.7 kDa protein.
GN P14G24.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federapfel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C.C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,

```

RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full Length cDNA of gene F14G24.20 (GI:12324636).";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Iehida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full Length cDNA of gene F14G24.20 (GI:12324636).";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AC019018; AAG52272.1; -;
 DR ENBL; AF370290; AAK44105.1; -;
 DR ENBL; AY063038; AAL34212.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 320 AA; 36696 MW; A686F44FCC592B3C CRC64;

 Query Match 92.3%; Score 36; DB 10; Length 320;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 VPNESE 7
 Db |||||
 222 VPNESE 228

 RESULT 5
 Q69071 PRELIMINARY; PRT; 1344 AA.
 AC Q69071;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DT Major capsid protein.
 OS Human herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=10372;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MUK;
 RX MEDLINE=96082321; PubMed=7483822;
 RA Mukai T., Isegawa Y., Yamaniishi K.;
 RT "Identification of the major capsid protein gene of human herpesvirus 7.";
 RL Virus Res. 37:55-62(1995).
 DR ENBL; D32005; BAA08770.1; -;
 DR InterPro; IPR000912; Herpes MCP.
 DR Pfam; PF03122; Herpes MCP; 1.
 DR PRINTS; PR00235; HSVCA5PSIDMCP.
 SQ SEQUENCE 1344 AA; 152782 MW; 8819A6DD801BA07E CRC64;

 Query Match 89.7%; Score 35; DB 12; Length 1344;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 PHNESE 7
 Db |||||
 1180 PHNESE 1185

 RESULT 6
 Q8W4H4 PRELIMINARY; PRT; 293 AA.
 ID Q8W4H4
 AC Q8W4H4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 26S proteasome regulatory subunit.
 GN AT2G39990.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AY062556; AAL32634.1; -;
 DR ENBL; AY093368; AAM13367.1; -;
 DR InterPro; IPR000555; Mov34.
 DR InterPro; IPR003640; Mov34_2.
 DR Pfam; PF01398; Mov34; 1.
 DR ProDom; PD005425; Mov34_2; 1.
 DR SMART; SM00232; JAB_MEN; 1.
 KW Proteasome.
 SQ SEQUENCE 293 AA; 31861 MW; 494C2F9E8F6F47C7 CRC64;

 Query Match 87.2%; Score 34; DB 10; Length 293;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 VPNESE 6
 Db |||||
 74 VPNESE 79

 RESULT 7
 O04202 PRELIMINARY; PRT; 293 AA.
 AC O04202;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 26S proteasome regulatory subunit S12 ISOLOG (Putative 26S proteasome regulatory subunit).
 GN T29M21.15 OR AT2G39990.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Rounsley S.D., Ketchum K.A., Lin X., Phillips C.A., Brandon R.C.,
 RA Fuhrmann J.L., White O., Kerlavage A.R., Adams M.D., Somerville C.R.,
 RA Venter J.C.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Asano K., Vornlocher H.-P., Richter-Cook N.J., Merrick W.C.,
 RA Hinnebusch A.G., Hershey J.W.B.;
 RL J. Biol. Chem. 0:0-0(1997).

[3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
 RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
 RA Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.,
 RA "Full Length cDNA of Gene T28M21.15/At2g39990 (GI:2088652).";
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.,
 RA "Arabidopsis Open Reading Frame (ORF) Clones";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RL ENBL; AF002109; AAB95284.1;
 DR ENBL; U54561; AAD03463.1;
 DR ENBL; AY045824; AAK76498.1;
 DR ENBL; AY091363; AAM14302.1;
 DR InterPro; IPR000555; Mov34;
 DR InterPro; IPR003640; Mov34_2;
 DR Pfam; PF01398; Mov34; 1.
 DR ProDom; PD005425; Mov34_2; 1.
 DR SMART; SM00232; JAB_MPN; 1.
 KW Initiation factor; Proteasome.
 SQ SEQUENCE 293 AA; 31862 MW; 4742CF7E8F6F47C7 CRC64;

Query Match 87.2%; Score 34; DB 10; Length 293;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
 |||||
 DB 74 VPHNES 79

RESULT 8
 O43060
 ID O43060 PRELIMINARY; PRT; 302 AA.
 AC O43060;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Hypothetical 33.3 kDa protein C4C3.07 in chromosome II.
 GN SPBC4C3.07.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,
 RA Duesterhoeft A.; (1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO PROTEASOME REGULATORY SUBUNIT S12/MOV-34.
 DR ENBL; AL021730; CAA16829.1;
 DR InterPro; IPR005555; Mov34.
 DR InterPro; IPR003640; Mov34_2.
 DR Pfam; PF01398; Mov34; 1.
 DR ProDom; PD005425; Mov34_2; 1.
 DR SMART; SM00232; JAB_MPN; 1.
 KW Hypothetical protein.

SQ SEQUENCE 302 AA; 33251 MW; A046E087CF083D84 CRC64;
 Query Match 87.2%; Score 34; DB 3; Length 302;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPHNES 6
 |||||
 DB 70 VPHNES 75
 RESULT 9
 Q921F7
 ID Q921F7 PRELIMINARY; PRT; 377 AA.
 AC Q921F7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to farnesyltransferase, CAAAX box, alpha.
 GN FNFA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR ENBL; BC012711; AAH12711.1;
 DR MGD; MGI:104683; Fnta.
 DR InterPro; IPR002088; PPTA.
 DR Pfam; PF01239; PPTA; 5.
 DR PROSITE; PS00904; PPTA; UNKNOWN_5.
 KW Transferase.
 SQ SEQUENCE 377 AA; 44012 MW; F3D8EF729D1D898C CRC64;
 Query Match 87.2%; Score 34; DB 11; Length 377;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPHNES 6
 |||||
 DB 270 VPHNES 275
 RESULT 10
 Q8U8U0
 ID Q8U8U0 PRELIMINARY; PRT; 381 AA.
 AC Q8U8U0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 8-amino-7-oxononanoate synthase.
 GN BIOF OR ATU3998 OR AGR_L1707.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.P. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";


```

RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Liu F.,
RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Leppas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58";
RL Science 294:2323-2328 (2001).
DR ENBL; AEO09331; AAL4800.1; -.
DR ENBL; AEO08283; AAK89427.1; -.
KW Complete proteome.
SQ SEQUENCE 381 AA; 39773 MW; 952964CB23BCE3FB CRC64;

Query Match 84.6%; Score 33; DB 16; Length 381;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
Db 140 VPHNEVE 146

RESULT 11
O97251 PRELIMINARY; PRT; 510 AA.
AC O97251;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative homologue of human EB1 protein.
GN PFC0305W, MAL3P2.30.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jaseal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrall B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum";
RL Nature 400:532-538 (1999).
DR ENBL; ALO34558; CAB39017.1; -.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PFO3271; EB1; 1.
DR PROSITE; PS50021; CH; 1.
SQ SEQUENCE 510 AA; 59544 MW; 29B08D620359325B CRC64;

Query Match 84.6%; Score 33; DB 5; Length 510;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 489 IPHNS 494

RESULT 12
Q9VY53 PRELIMINARY; PRT; 115 AA.
ID Q9VY53
AC Q9VY53;

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG12482 protein.
GN CG11595 OR CG12482.
OS Drosophila melanogaster (Fruit Fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neozoa; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svendsen R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
DR ENBL; AEO03495; AAF48352.1; -.
DR FlyBase; FBgn0030548; CG11595.
SQ SEQUENCE 115 AA; 13640 MW; CEE29A8C8BF402A8 CRC64;

Query Match 82.1%; Score 32; DB 5; Length 115;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
Db 56 VPHNEKD 62

RESULT 13
Q9LES7 PRELIMINARY; PRT; 332 AA.
ID Q9LES7
AC Q9LES7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 37.1 kDa protein.

```

GN T8M16 140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benes V., Wurnbach E., Drzonek H., Ansoerge W., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X., Quetier P., Salanoubat M.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; ALJ90921; CAC00744.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 332 AA; 37094 MW; F68FA19855BB956E CRC64;

Query Match 82.1%; Score 32; DB 10; Length 332;

Best Local Similarity 71.4%; Pred. No. 96;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
 :|||||
 Db 88 LPNHEGE 94

RESULT 14

O93345
 ID O93345 PRELIMINARY; PRT; 369 AA.
 AC O93345;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Transcription factor LEP-1.
 GN LEP-1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98259088; PubMed=9596593;
 RA Kengaku M., Capdevilla J., Rodriguez-Esteban C., De La Pena J.,
 RA Johnson R.L., Belmonte J.C.I., Tabin C.J.;
 RT "Distinct WNT pathways regulating AER formation and dorsoventral
 RT polarity in the chick limb bud."
 RL Science 280:1274-1277(1998).
 DR EMBL; AF064462; AAC24524.1; -;
 DR HSSP; P27782; ZLEP.
 DR InterPro; IPR000910; HMG_12_box.
 DR Pfam; PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 SQ SEQUENCE 369 AA; 40857 MW; BADF4B097103F864 CRC64;

Query Match 82.1%; Score 32; DB 13; Length 369;

Best Local Similarity 83.3%; Pred. No. 11e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
 :|||||
 Db 243 PHNDSE 248

RESULT 15

Q05699
 ID Q05699 PRELIMINARY; PRT; 372 AA.
 AC Q05699;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE ASF-1/G13 protein.
 GN ASF-1/G13
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017650; PubMed=1921969;
 RA Fromm H., Katagiri F., Chua N.-H.;
 RT "The tobacco transcription activator TGA1 binds to a sequence in the
 RT 5' upstream region of a gene encoding a TGA1-related protein.";
 RL Mol. Gen. Genet. 229:181-188(1991).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
 DR EMBL; M62855; AAA34091.1; -;
 DR TRANSFAC; T04856; -;
 DR InterPro; IPR004827; TF_bZIP.
 DR Pfam; PF00170; bZIP; 1.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS00036; BZIP_BASIC; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 372 AA; 41991 MW; 56A2730610C9A06E CRC64;

Query Match 82.1%; Score 32; DB 10; Length 372;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
 :|||||
 Db 60 PHNETE 65

Search completed: May 6, 2003, 15:02:51
 Job time: 27.85 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:58:24 ; Search time 7 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: US-09-851-422B-8

Perfect score: 39
Sequence: 1 VPHNESE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	357	1 IF35 HUMAN	O00303 homo sapien
2	35	89.7	1345	1 VCAP_HSV7J	P52347 human herpe
3	34	87.2	340	1 PFTA BOVIN	P29702 bos taurus
4	34	87.2	377	1 PFTA MOUSE	O61239 mus musculu
5	34	87.2	377	1 PFTA RAT	O04631 rattus norv
6	34	87.2	379	1 PFTA HUMAN	P49354 homo sapien
7	32	82.1	688	1 YAOA_SCHPO	Q10109 schizosacch
8	31	79.5	182	1 RL18_HALNI	P50562 halobacteri
9	31	79.5	254	1 HM36 CAEEL	Q93352 caenorhabdi
10	31	79.5	316	1 CORA ECOLI	P27841 escherichia
11	31	79.5	316	1 CORA SALTU	P31138 salmonella
12	31	79.5	316	1 Y275_CHLMU	O9pl33 chlamydia m
13	31	79.5	2156	1 RPI1 HUMAN	P56715 homo sapien
14	30	76.9	278	1 CEMA_CYACA	O9tm16 cyanidum c
15	30	76.9	315	1 CORA HAEIN	P44998 haemophilus
16	30	76.9	404	1 CTNS CAEEL	Q09500 caenorhabdi
17	30	76.9	409	1 THBG RAT	P35577 rattus norv
18	30	76.9	421	1 SYTL1 MOUSE	P46096 mus musculu
19	30	76.9	421	1 SYTL1 RAT	P21707 rattus norv
20	30	76.9	461	1 PRTM HUMAN	P04070 homo sapien
21	30	76.9	523	1 YQE4 CAEEL	Q17529 caenorhabdi
22	30	76.9	752	1 8511 TRYCR	P18269 trypanosoma
23	30	76.9	827	1 REDI1 YEAST	P14291 saccharomyc
24	30	76.9	901	1 PODK TREPA	O83728 treponema p
25	30	76.9	982	1 N1A_NEUCR	P08619 neurospora
26	30	76.9	1435	1 Y194 HUMAN	Q12766 homo sapien
27	30	76.9	1448	1 PK3G HUMAN	O75747 homo sapien
28	29	74.4	244	1 PYRF_XYLFA	O9pbh0 xylella fas
29	29	74.4	271	1 RL5_DUNSA	O22608 dunaliella
30	29	74.4	293	1 RL5_CAEL	P49405 caenorhabdi
31	29	74.4	319	1 YIBQ ECOLI	P37891 escherichia
32	29	74.4	355	1 Y497 MYCLE	P54580 mycobacteri
33	29	74.4	406	1 YG43 YEAST	P53298 saccharomyc

Query Match 100.0% ; Score 39 ; DB 1 ; Length 357 ;
Best Local Similarity 100.0% ; Pred. No. 0.89 ;
Matches 7 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

ALIGNMENTS

RESULT 1

IF35_HUMAN STANDARD; PRT; 357 AA.
AC O00303;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 5 (eIF-3 epsilon)
DE (eIF3 p47 subunit).
GN EIF3S5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98001678; PubMed=9341143;
RA Asano K., Vornlocher H.-P., Richter-Cook N.J., Merrick W.C.,
RA Hinnebusch A.G., Hershey J.W.B.;
RT "Structure of cDNAs encoding human eukaryotic initiation factor 3
RT subunits. Possible roles in RNA binding and macromolecular assembly." ;
RL J. Biol. Chem. 272:27042-27052(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
CC METHIONYL-TRNAI AND MRNA. ASSOCIATES WITH THE COMPLEX P170-EIF3.
CC -!- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE PSMD7/C6.1A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U94855; AAD03467.1; -;
DR EMBL; BC000490; AAH00490.1; -;
DR Genew; HGNC:3275; EIF3S5.
DR MIM; 603914; -;
DR InterPro; IPR000555; Mov34.
DR InterPro; IPR003640; Mov34_2.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD005425; Mov34_2; 1.
DR SMART; SM00232; JAB MPN; 1.
KW Initiation factor; Protein biosynthesis.
SQ SEQUENCE 357 AA; 37564 MW; 8A70FC6E2BF07737 CRC64;

Q03750 saccharomyc
Q23695 crithidia f
P17054 rhodobacter
Q9dba9 mus musculu
P32780 homo sapien
Q64355 mus musculu
P19880 saccharomyc
O10282 orgyia pseu
P16788 human cytom
Q45409 ralestonia s
P58593 ralestonia s
P53394 saccharomyc

34 29 74.4 510 1 YML4 YEAST
35 29 74.4 515 1 DRTS CRIFA
36 29 74.4 524 1 CRTI RHOCA
37 29 74.4 547 1 TFI1 MOUSE
38 29 74.4 548 1 TFI1 HUMAN
39 29 74.4 560 1 EFS MOUSE
40 29 74.4 650 1 PDR4 YEAST
41 29 74.4 657 1 Y023 NPVOP
42 29 74.4 707 1 GCVK HCMVA
43 29 74.4 750 1 EPB2 RALSO
44 29 74.4 751 1 EPB1 RALSO
45 29 74.4 754 1 SULX YEAST

QY 1 VPNESE 7
DB 137 VPNESE 143
|||||

RESULT 2
VCAP HSV7J STANDARD; PRT; 1345 AA.
AC P52347;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP).
GN US7.
OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U43400; AAC54720.1; -;
DR InterPro; IPR000912; Herpes MCP.
DR Pfam; PF03122; Herpes MCP; I.
DR PRINTS; PR00235; HSVCAPSIDMCP.
KW Coat protein.
SQ SEQUENCE 1345 AA; 152931 MW; 93F4D3FF9BF4977E CRC64;
Query Match 89.7%; Score 35; DB 1; Length 1345;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
DB 1181 PHNESE 1186
|||||

RESULT 3
ID PFTA_BOVIN STANDARD; PRT; 340 AA.
AC P29702;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX farnesyltransferase alpha subunit) (RAS proteins prenyltransferase alpha) (Ftase-alpha) (Fragment).
GN FNTA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92011456; PubMed=1918005;
RA Kohl N.B.; Diehl R.E.; Schaber M.D.; Rands E.; Soderman D.D.; He B.;
RA Moores S.L.; Pompliano D.L.; Ferro-Novick S.; Powers S.; Thomas K.A.;
RA Gibbs J.B.;

RT "Structural homology among mammalian and Saccharomyces cerevisiae isoprenyl-protein transferases";
RL J. Biol. Chem. 266:18884-18888(1991).
CC -!- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIETY FROM FARNESYL PYROPHOSPHATE TO A CYSTEINE AT THE FOURTH POSITION FROM THE C-TERMINUS OF SEVERAL PROTEINS. THE ALPHA SUBUNIT IS THOUGHT TO PARTICIPATE IN A STABLE COMPLEX WITH THE SUBSTRATE FARNESYL-PP.
CC THE BETA SUBUNIT BINDS THE PEPTIDE SUBSTRATE.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE ALPHA SUBUNIT FAMILY.
CC
CC -!- SIMILARITY: CONTAINS 5 PFTA REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M74083; AAA30529.1; ALT_INIT.
DR PIR; A41013; A41013.
DR HSSP; Q04631; 1FT1.
DR InterPro; IPR002088; PPTA.
DR Pfam; PF01239; PPTA; 5.
DR PROSITE; PS00904; PPTA; 5.
KW Transferase; Prenyltransferase; Repeat.
FT NON_TER 1
FT REPEAT 76 109 PPTA 1.
FT REPEAT 111 144 PPTA 2.
FT REPEAT 145 178 PPTA 3.
FT REPEAT 180 212 PPTA 4.
FT REPEAT 219 252 PPTA 5.
SQ SEQUENCE 340 AA; 40494 MW; F884B68203D24E6F CRC64;
Query Match 87.2%; Score 34; DB 1; Length 340;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPNESE 6
DB 231 VPNESE 236
|||||

RESULT 4
ID PFTA_MOUSE STANDARD; PRT; 377 AA.
AC Q61239;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX farnesyltransferase alpha subunit) (RAS proteins prenyltransferase alpha) (Ftase-alpha).
GN FNTA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96069614; PubMed=7590362;
RA Shirasawa H.; Kinoshita T.; Shino Y.; Mori K.; Shimizu K.;
RA Simizu B.;
RT "Cloning and sequencing of the murine farnesyltransferase alpha-encoding cDNA from a cell line which expresses the human papillomavirus type-16 E6 gene."
RL Gene 154:373-374(1995).
CC -!- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIETY FROM FARNESYL PYROPHOSPHATE TO A CYSTEINE AT THE FOURTH POSITION FROM THE C-TERMINUS OF SEVERAL PROTEINS. THE ALPHA SUBUNIT IS THOUGHT TO PARTICIPATE IN A STABLE COMPLEX WITH THE SUBSTRATE FARNESYL-PP.
CC

CC THE BETA SUBUNIT BINDS THE PEPTIDE SUBSTRATE.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIETY FROM
 CC FARNESYL PYROPHOSPHATE TO A CYSTEINE AT THE FOURTH POSITION FROM
 CC THE C-TERMINUS OF SEVERAL PROTEINS. THE ALPHA SUBUNIT IS THOUGHT
 CC TO PARTICIPATE IN A STABLE COMPLEX WITH THE SUBSTRATE FARNESYL-PP.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D49744; BAA08578.1; -
 CC HSSP; Q04631; 1PT1.
 CC MGD; MGI:104683; PPTA.
 CC InterPro: IPR002088; PPTA.
 CC Pfam; PF01239; PPTA; 5.
 CC PROSITE; PS00904; PPTA; 5.
 CC TRANSFERASE; Prenyltransferase; Repeat.
 CC DOMAIN 22 30 POLY-PRO.
 CC REPEAT 115 148 PPTA 1.
 CC REPEAT 150 183 PPTA 2.
 CC REPEAT 184 217 PPTA 3.
 CC REPEAT 219 251 PPTA 4.
 CC REPEAT 258 291 PPTA 5.
 CC SEQUENCE 377 AA; 44013 MW; F3D60B9899F36D66 CRC64;
 SQ
 Query Match 87.2%; Score 34; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPHNES 6
 |||||
 DB 270 VPHNES 275
 RESULT 5
 PPTA_RAT STANDARD; PRT; 377 AA.
 AC Q04631;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX
 DE farnesyltransferase alpha subunit) (RAS proteins prenyltransferase
 DE alpha) (FTase-alpha).
 GN FNTA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92107951; PubMed=1763049;
 RA Chen W.-J., Andres D.A., Goldstein J.L., Brown M.S.;
 RT "Cloning and expression of a cDNA encoding the alpha subunit of rat
 RT p13ras protein farnesyltransferase."
 RL Proc. Natl. Acad. Sci. U.S.A. 88:11368-11372 (1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).
 RX MEDLINE=97218306; PubMed=9065406;
 RA Park H.-W., Boduluri S.R., Moomaw J.F., Casey P.J., Beese L.S.;
 RT "Crystal structure of protein farnesyltransferase at 2.25-A
 RT resolution."
 RL Science 275:1800-1804 (1997).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).
 RX MEDLINE=98322062; PubMed=9657673;
 RA Long S.B., Casey P.J., Beese L.S.;
 RT "Cocrystal structure of protein farnesyltransferase complexed with a

RT farnesyl diphosphate substrate.";
 RL Biochemistry 37:9612-9618 (1998).
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIETY FROM
 CC FARNESYL PYROPHOSPHATE TO A CYSTEINE AT THE FOURTH POSITION FROM
 CC THE C-TERMINUS OF SEVERAL PROTEINS. THE ALPHA SUBUNIT IS THOUGHT
 CC TO PARTICIPATE IN A STABLE COMPLEX WITH THE SUBSTRATE FARNESYL-PP.
 CC
 CC THE BETA SUBUNIT BINDS THE PEPTIDE SUBSTRATE.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE ALPHA SUBUNIT
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 PPTA REPEATS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M81225; AAA41833.1; -
 CC PIR; A41625; A41625.
 CC PDB; 1FT1; 18-MAR-98.
 CC PDB; 1FT2; 18-NOV-98.
 CC InterPro: IPR002088; PPTA.
 CC Pfam; PF01239; PPTA; 5.
 CC PROSITE; PS00904; PPTA; 5.
 CC TRANSFERASE; Prenyltransferase; Repeat; 3D-structure.
 CC DOMAIN 22 30 PRO-RICH.
 CC REPEAT 115 148 PPTA 1.
 CC REPEAT 150 183 PPTA 2.
 CC REPEAT 184 217 PPTA 3.
 CC REPEAT 219 251 PPTA 4.
 CC REPEAT 258 291 PPTA 5.
 CC SEQUENCE 377 AA; 44049 MW; DFFECCLB88BC080 CRC64;
 SQ
 Query Match 87.2%; Score 34; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPHNES 6
 |||||
 DB 270 VPHNES 275
 RESULT 6
 PPTA_HUMAN STANDARD; PRT; 379 AA.
 AC P49354; Q9UDC1;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX
 DE farnesyltransferase alpha subunit) (RAS proteins prenyltransferase
 DE alpha) (FTase-alpha).
 GN FNTA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=94102736; PubMed=8276393;
 RA Andres D.A., Milatovich A., Ozelik T., Wenzlau J.M., Brown M.S.,
 RA Goldstein J.L., Francke U.;
 RT "cDNA cloning of the two subunits of human CAAX farnesyltransferase
 RT and chromosomal mapping of FNTA and FNTB loci and related
 RT sequences".
 RL Genomics 18:105-112 (1993).
 RN [2]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.
 RX MEDLINE=93123261; PubMed=8419339;

RA Andres D.A., Goldstein J.L., Ho Y.K., Brown M.S.;
 RT "Mutational analysis of alpha-subunit of protein farnesyltransferase.
 Evidence for a catalytic role";
 RL J. Biol. Chem. 268:1383-1390(1993).
 [3]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Colon;
 RC Submitted to the EMBL/GenBank/DBJ databases.
 RL Sraueberg R.;
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIEITY FROM
 CC FARNESYL PYROPHOSPHATE TO A CYSTEINE AT THE FOURTH POSITION FROM
 CC THE C-TERMINUS OF SEVERAL PROTEINS. THE ALPHA SUBUNIT IS THOUGHT
 CC TO PARTICIPATE IN A STABLE COMPLEX WITH THE SUBSTRATE FARNESYL-PP.
 CC THE BETA SUBUNIT BINDS OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE ALPHA SUBUNIT
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 PPTA REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; L10413; AAA6285.1; -;
 DR ENBL; L00634; AAA3583.1; -;
 DR ENBL; BC017029; AAH17029.1; -;
 DR HSP; Q04631; IPT1.
 DR Genew; HGNC:3782; FNTA.
 DR MTM; I34635; -;
 DR InterPro; IPR002088; PPTA.
 DR Pfam; PF01239; PPTA; 5.
 DR PROSITE; PS00904; PPTA; 5.
 DR Transferase; Prenyltransferase; Repeat.
 KW DOMAIN 22 31 PRO-RICH.
 FT REPEAT 115 148 PPTA 1.
 FT REPEAT 150 183 PPTA 2.
 FT REPEAT 184 217 PPTA 3.
 FT REPEAT 219 251 PPTA 4.
 FT REPEAT 258 291 PPTA 5.
 FT MUTAGEN 164 164 K->N: REDUCED ACTIVITY.
 FT CONFLICT 241 241 Y -> H (IN REF. 2).
 SQ SEQUENCE 379 AA; 44408 MW; E933CBA874AB92B9 CRC64;
 Query Match 87.2%; Score 34; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPHNES 6
 DB 270 VPHNES 275
 RESULT 7
 ID YAOA SCHPO STANDARD; PRT; 688 AA.
 AC Q10109;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C18G6.10 in chromosome 1.
 GN SPAC18G6.10.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver P., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welljens J., Vanacker E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Purnelle B.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armatrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; Z68198; CAA92388.1; -;
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 547 587 POTENTIAL.
 SQ SEQUENCE 688 AA; 78178 MW; 3CB29ED7B7A5DEA2 CRC64;
 Query Match 82.1%; Score 32; DB 1; Length 688;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PHNESE 7
 DB 97 PHNETE 102
 RESULT 8
 ID RL18 HALN1 STANDARD; PRT; 182 AA.
 AC P50562; Q9HPB5;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L18P (HSA118).
 GN RPL18P OR VNG1714G.
 OS Halobacterium sp. (strain NRC-1), and
 OS Halobacterium salinarum.
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091, 2242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRC-1;
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE OF 1-23.
RC SPECIES-H.salinarius; STRAIN=DSM 3754;
RX MEDLINE=94229075; PubMed=8174557;
RA McDougall J., Wittmann-Liebold B.;
RT "Comparative analysis of the protein components from 5S rRNA protein
RT complexes of halophilic archaeobacteria.";
RL Eur. J. Biochem. 221:779-785(1994).
CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A5005077; AG19955.1; -.
DR FIR; D33084; D33084.
DR InterPro; IPR001149; Ribosomal_L18p.
DR Pfam; PF00861; Ribosomal_L18p; 1.
KW Ribosomal protein; Complete proteome.
FT INIT MET 0
SQ SEQUENCE 182 AA; 19680 MW; 1607FD9C5243288B CRC64;

Query Match 79.5%; Score 31; DB 1; Length 182;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPINES 6
Db 129 VPHNDS 134
|||||:

RESULT 9
HM36_CABEL STANDARD; PRT; 254 AA.
AC Q93352;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-MAY-2000 (Rel. 39, Last annotation update)
DE Homeobox protein ceh-36.
GN CEH-36 OR C37E2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Spring J.;
RL Unpublished observations (JUN-1999).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE PROPOSED SPLICING
CC PATTERN HAS BEEN REVISED TO PRODUCE A CORRECT PAIRED-TYPE HOMEBOX
CC REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z81046; CAB02821.1; ALT_SEQ.
DR HSSP; P06601; 1FJL.
DR TRANSFAC; T03364;
DR WormPep; C37E2.4; CE08623.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF0046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA BIND 55 111 HOMEBOX.
FT DOMAIN 171 182 POLY-GLN.
SQ SEQUENCE 254 AA; 28268 MW; 2720CB47A45660A7 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 254;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PHNESE 7
Db 133 PHNESK 138
|||||:

RESULT 10
COR_A_ECOLI STANDARD; PRT; 316 AA.
AC P27841;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Magnesium and cobalt transport protein corA.
GN CORA OR B3816 OR 25333 OR ECS4746.
OS Escherichia coli, and
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93300795; PubMed=8314774;
RA Smith R.L., Banks J.L., Snavely M.D., Maguire M.E.;
RT "Sequence and topology of the CorA magnesium transport systems of
RT Salmonella typhimurium and Escherichia coli. Identification of a new
RT class of transport protein.";
RL J. Biol. Chem. 268:14071-14080(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Ohmori H.;
RL Submitted (XXY-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
RN [4]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;


```

ID Y275 CHLMU STANDARD; PRT; 316 AA.
AC Q9PL33;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TC0275.
GN TC0275.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni99;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0441/CT007/TC0275
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A5002295; AAP39143.1; -
CC TIGR; TC0275; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 316 AA; 35475 MW; 0E10E52452CC9A66 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 316;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 PHNESE 7
Db 292 PHNEDE 297
|||||
292 PHNEDE 297

RESULT 13
RPI_HUMAN
ID RPI_HUMAN STANDARD; PRT; 2156 AA.
AC P56715;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxygen-regulated protein 1 (Retinitis pigmentosa RPI protein)
DE (Retinitis pigmentosa 1 protein).
GN RPI OR ORP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS HIS-872; TYR-985; THR-1670; PRO-1691
RP AND TYR-2033.
RC TISSUE=Retina;
RX MEDLINE=99318096; PubMed=10391211;
RA Sullivan L.S., Heckenlively J.R., Bowne S.J., Zuo J., Hide W.A.,
RA Gail A., Denton M., Inglehearn C.F., Blanton S.H., Daiger S.P.;
RT "Mutations in a novel retina-specific gene cause autosomal dominant
RT retinitis pigmentosa."
RL Nat. Genet. 22:255-259(1999).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Retina;
RX MEDLINE=99318096; PubMed=10391211;
RA Pierce E.A., Quinn T., Meehan T., McGee T.L., Berson E.L., Dryja T.P.;
RT "Mutations in a gene encoding a new oxygen-regulated photoreceptor
RT protein cause dominant retinitis pigmentosa."
RL Nat. Genet. 22:248-254(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99330563; PubMed=10401003;
RA Guillemeau X., Piriev N.I., Danciger M., Kozak C.A., Cideciyan A.V.,
RA Jacobson S.G., Farber D.B.;
RT "A nonsense mutation in a novel gene is associated with retinitis
RT pigmentosa in a family linked to the RPL locus."
RL Hum. Mol. Genet. 8:1541-1546(1999).
CC -!- FUNCTION: COULD HAVE A ROLE IN THE DIFFERENTIATION OF
CC PHOTORECEPTOR CELLS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN RETINA. NOT EXPRESSED IN HEART,
CC BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, SPLEEN AND
CC PANCREAS.
CC -!- DISEASE: DEFECTS IN RPL CAUSE RETINITIS PIGMENTOSA FORM 1 (RPL); A
CC DISEASE CHARACTERIZED BY CONSTRUCTION OF THE VISUAL FIELDS, NIGHT
CC BLINDNESS, AND FUNDUS CHANGES. THE DISEASE SEEMS TO BE ASSOCIATED
CC WITH TRUNCATED (STOP OR FRAMESHIFT MUTATIONS) FORMS OF THE
CC PROTEIN.
CC -!- SIMILARITY: CONTAINS 2 DOUBLECORTIN DOMAINS.
CC -!- DATABASE: NAME=RetNet;
CC NOTE=Retinal information network;
CC WWW="http://www.sph.uth.tmc.edu/tetnet/".
CC -!- DATABASE: NAME=Mutations of the RPL gene;
CC NOTE=Retina International's Scientific Newsletter;
CC WWW="http://www.retina-international.com/sci-news/rplmut.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF143226; AAD44197.1; -
CC EMBL; AF143224; AAD44197.1; JOINED.
CC EMBL; AF143225; AAD44197.1; JOINED.
CC EMBL; AF143222; AAD44198.1; -
CC EMBL; AF141021; AAD42072.1; -
CC EMBL; AF152242; AAD46774.1; -
CC EMBL; AF152240; AAD46774.1; JOINED.
CC EMBL; AF152241; AAD46774.1; JOINED.
CC EMBL; AF146592; AAD46769.1; -
CC GENE; HGNC:10263; RPL.
CC MIM; 603937; -
CC MIM; 180100; -
CC InterPro; IPR003533; DCX.
CC Pfam; PF03607; DCX; 2.
CC SMART; SM00537; DCX; 2.
CC PROSITE; PS03039; DC; 2.
KW Vision; Retinitis pigmentosa; Polymorphism; Repeat.
FT DOMAIN 36 118 DOUBLECORTIN 1.
FT DOMAIN 154 233 DOUBLECORTIN 2.
FT DOMAIN 268 273 POLY-SER.
FT DOMAIN 671 675 POLY-LYS.
FT DOMAIN 1687 1691 POLY-SER.
FT VARIANT 872 872 R -> H.
FT VARIANT 985 985 N -> Y.
FT VARIANT 1670 1670 A -> T.
FT VARIANT 1691 1691 S -> P.
FT VARIANT 2033 2033 C -> Y.
FT SEQUENCE 2156 AA; 240659 MW; 55AEDBEC43D6A507 CRC64;
SQ

```

Query Match 79.5%; Score 31; DB 1; Length 2156;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
 |||||
 DB 803 PHNESK 808

RESULT 14
 CEMA CYACA STANDARD; PRT; 278 AA.
 AC Q9TW16;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chloroplast envelope membrane protein.
 GN CEMA OR YCF10 OR YCF56.
 OS Cyanidium caldarium.
 OG Chloroplast.

OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
 OC Cyanidium.
 OX NCBI_TaxID=2771;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=RK-1;
 RX MEDLINE=20496959; PubMed=11040290;
 RA Gloeckner G., Rosenthal A., Valentin K.-U.;
 RT "The structure and gene repertoire of an ancient red algal plastid
 genome.";

RL J. Mol. Evol. 51:382-390(2000).
 CC -!- FUNCTION: May be involved in proton extrusion. Indirectly promotes
 CC efficient inorganic carbon uptake into chloroplasts (by
 CC similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast inner
 CC envelope (by similarity).
 CC -!- SIMILARITY: BELONGS TO THE CEMA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF022186; AAF12995.1; -
 DR InterPro; IPR004282; Cema.
 DR Pfam; PF03040; Cema; 1.
 KW Chloroplast; Transmembrane; Transport; Hydrogen ion transport.
 FT TRANSMEM 61 81 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 203 223 POTENTIAL.
 FT TRANSMEM 239 259 POTENTIAL.
 SQ SEQUENCE 278 AA; 32355 MW; 88B1050D7836102E CRC64;

Query Match 76.9%; Score 30; DB 1; Length 278;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
 |||||
 DB 37 PHSESE 42

RESULT 15
 CORA_HAEIN STANDARD; PRT; 315 AA.
 AC P44998;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Magnesium and cobalt transport protein corA.
 CORA OR H11035.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.P., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: PLAYS A ROLE IN THE TRANSPORT OF MAGNESIUM AND COBALT
 CC IONS (BY SIMILARITY).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; U32784; AAC22695.1; -
 DR TIGR; H11035; -
 DR InterPro; IPR002523; CorA.
 DR InterPro; IPR004488; CorA_transp.
 DR Pfam; PF01544; CorA; 1.
 DR TIGRfam; TIGR00383; corA; 1.
 KW Magnesium; Cobalt; Transmembrane; Transport; Complete proteome.
 FT TRANSMEM 257 277 POTENTIAL.
 FT TRANSMEM 289 309 POTENTIAL.
 SQ SEQUENCE 315 AA; 36593 MW; DA4EDA284CC68DCC CRC64;

Query Match 76.9%; Score 30; DB 1; Length 315;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNES 6
 |||||
 DB 227 PHNES 231

Search completed: May 6, 2003, 15:01:31
 Job time : 9 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:05 ; Search time 12.25 Seconds
(without alignments)
54.934 Million cell updates/sec

Title: US-09-851-422B-8
Perfect score: 39
Sequence: 1 VPHNESE 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	92.3	320	2 F96570	unknown protein, 8
2	35	89.7	1345	2 T41960	major capsid prote
3	34	87.2	293	2 H84823	26S proteasome reg
4	34	87.2	302	2 T40490	probable 26S prote
5	34	87.2	340	2 A41013	protein farnesyltr
6	34	87.2	377	2 JC4368	protein farnesyltr
7	34	87.2	377	2 A41825	protein farnesyltr
8	34	87.2	379	2 A47659	farnesyl-protein t
9	33	84.6	381	2 A96238	bioF protein (AF31
10	33	84.6	381	2 AB3048	8-amino-7-oxonon
11	32	82.1	332	2 T51269	hypothetical prote
12	32	82.1	372	2 S17715	transcription acti
13	32	82.1	688	2 T37923	hypothetical prote
14	32	82.1	1819	2 D97132	uncharacterized ph
15	31	79.5	183	2 G84323	50S ribosomal prot
16	31	79.5	249	2 D96891	hypothetical prote
17	31	79.5	257	2 T19809	hypothetical prote
18	31	79.5	272	2 T26235	hypothetical prote
19	31	79.5	288	2 T48753	conserved hypotet
20	31	79.5	316	2 A47157	magnesium transpor
21	31	79.5	316	2 B47157	magnesium transpor
22	31	79.5	316	2 AE0467	magnesium and coba
23	31	79.5	316	2 AF0918	magnesium and coba
24	31	79.5	316	2 B86068	Mg2+ transport, sy
25	31	79.5	316	2 B91222	Mg2+ transport pro
26	31	79.5	316	2 C81720	conserved hypotet
27	31	79.5	374	2 T15895	hypothetical prote
28	31	79.5	382	2 T25538	hypothetical prote
29	31	79.5	548	2 T25424	hypothetical prote

30	31	79.5	1068	2 G86452	hypothetical prote
31	31	79.5	1467	2 T48162	hypothetical prote
32	30	76.9	97	2 T45416	hypothetical prote
33	30	76.9	143	2 T45444	hypothetical prote
34	30	76.9	172	2 PNO545	gamma4-crystallin
35	30	76.9	220	2 AE2360	hypothetical prote
36	30	76.9	266	2 H69468	lysophospholipase
37	30	76.9	315	2 A64109	magnesium transpor
38	30	76.9	319	2 A70638	hypothetical prote
39	30	76.9	360	2 G64575	conserved hypotet
40	30	76.9	382	2 E85082	hypothetical prote
41	30	76.9	382	2 T14186	hypothetical prote
42	30	76.9	387	2 G72288	conserved hypotet
43	30	76.9	418	2 A39567	thyroxine-binding
44	30	76.9	421	2 S09595	synaptotagmin P65
45	30	76.9	423	2 F96633	hypothetical prote

ALIGNMENTS

RESULT 1

F96570
unknown protein, 80333-82175 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96570
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96570
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <STO>
A:Cross-references: GB:AE005173; NID:G6862953; PIDN:AAF30341.1; GSPDB:GN00141
C:Genetics:
A:Gene: F14G24.20
A:Map position: 1
C:Superfamily: probable membrane protein YOLO77c

Query Match 92.3%; Score 36; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 6.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy' 1 VPHNESE 7

Db 222 VPHNESD 228

RESULT 2

T41960
major capsid protein - human herpesvirus 7 (strain J1)
C:Species: human herpesvirus 7
A:Variety: strain J1
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T41960
R:Nicholas, J.
submitted to the EMBL Data Library, December 1995
A:Description: Determination and analysis of the complete nucleotide sequence of human
A:Reference number: Z22022
A:Accession: T41960
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1345 <NIC>
A:Cross-references: EMBL:U43400; PIDN:AAC54720.1

A;Experimental source: strain J1

C;Genetics:

A;Note: U57

C;Superfamily: varicella-zoster virus major capsid protein

Query Match 89.7%; Score 35; DB 2; Length 1345;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7

|||||

Db 1181 PHNESE 1186

RESULT 3

H84823

26S proteasome regulatory subunit [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: H84823

R;Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: H84823

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-293 <STO>

A;Cross-references: GB:AE002093; NID:g2088652; PIDN:AAB95284.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g39990

A;Map position: 2

Query Match 87.2%; Score 34; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

|||||

Db 74 VPHNES 79

RESULT 4

T40490

Probable 26S proteasome regulatory subunit - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C;Accession: T40490

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.

submitted to the EMBL Data Library, February 1998

A;Reference number: Z21910

A;Accession: T40490

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-302 <WOO>

A;Cross-references: EMBL:AL021730; PIDN:CAAL6829.1; GSPDB:GN00067; SPDB:SPBC4C3.07

A;Experimental source: strain 972h-; cosmid c4C3

C;Genetics:

A;Gene: SPDB:SPBC4C3.07

A;Map position: 2

C;Superfamily: mov-34 protein

Query Match 87.2%; Score 34; DB 2; Length 302;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

|||||

Db 70 VPHNES 75

RESULT 5

A41013

protein farnesyltransferase (EC 2.5.1.1-) alpha chain - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 28-May-1993

C;Accession: A41013

R;Kohl, N.E.; Diehl, R.E.; Schaber, M.D.; Rands, E.; Soderman, D.D.; He, B.; Moores, S.;

J. Biol. Chem. 266, 18984-18988, 1991

A;Title: Structural homology among mammalian and Saccharomyces cerevisiae isoprenyl-pro

A;Reference number: A41013; MUID:92011656; PMID:1918005

A;Accession: A41013

A;Molecule type: mRNA

A;Residues: 1-340 <KOH>

A;Cross-references: GB:M74083

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: This protein attaches farnesyl residues to a cysteine near the carboxyl term

C;Keywords: heterodimer; transferase

Query Match 87.2%; Score 34; DB 2; Length 340;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

|||||

Db 231 VPHNES 236

RESULT 6

JC4368

protein farnesyltransferase (EC 2.5.1.1-) alpha chain - mouse

N;Alternate names: farnesyltransferase alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 10-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 20-Jun-2000

C;Accession: JC4368

R;Shirasawa, H.; Kinoshita, T.; Shino, Y.; Mori, K.; Shimizu, K.; Simizu, B.

Gene 164, 373-374, 1995

A;Title: Cloning and sequencing of the murine farnesyltransferase alpha-encoding cDNA fr

A;Reference number: JC4368; MUID:96069614; PMID:7590362

A;Accession: JC4368

A;Molecule type: mRNA

A;Residues: 1-377 <SHI>

A;Cross-references: DDBJ:D49744; NID:g1136744; PIDN:BAA08578.1; PID:g1136745

A;Experimental source: 10T1/2 cell line

C;Comment: This enzyme, which contains two subunits, alpha and beta, transfers the farn

g nuclear lamins and retinal proteins. This enzyme is a heterodimeric enzyme and alpha

C;Genetics:

A;Gene: ita

C;Keywords: heterodimer; transferase

Query Match 87.2%; Score 34; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

|||||

Db 270 VPHNES 275

RESULT 7

A41625

protein farnesyltransferase (EC 2.5.1.1-) alpha chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 05-Nov-1999

C;Accession: A41625

R;Chen, W.J.; Andres, D.A.; Goldstein, J.L.; Brown, M.S.

Proc. Natl. Acad. Sci. U.S.A. 88, 11368-11372, 1991

A;Title: Cloning and expression of a cDNA encoding the alpha subunit of rat p21(ras) pr

A;Reference number: A41625; MUID:92107951; PMID:1763049

A;Accession: A41625

A;Molecule type: mRNA

A;Residues: 1-377 <CHE>

A;Cross-references: GB:M81225; NID:g206093; PIDN:AAA41833.1; PID:g206094

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: This protein attaches farnesyl residues to a cysteine near the carboxyl termi
C;Keywords: heterodimer; transferase

Query Match 87.2%; Score 34; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
|||||

DB 270 VPHNES 275

RESULT 8
A47659
N;Alternate names: CAAX farnesyltransferase alpha chain - human
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 10-May-1996 #text_change 21-Jul-2000
C;Accession: A47659; A92274
R;Andres, D.A.; Milatovich, A.; Ozelik, T.; Wenzlau, J.M.; Brown, M.S.; Goldstein, J.L.
Genomics 18, 105-112, 1993
A;Title: cDNA cloning of the two subunits of human CAAX farnesyltransferase and chromoso
A;Reference number: A47659; MUID:94102736; PMID:8276393
A;Accession: A47659
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-379 <AND>
A;Cross-references: GB:L10413; NID:G388755; PIDN:AAA66285.1; PID:G388756
R;Omer, C.A.; Kral, A.M.; Diehl, R.E.; Prendergast, G.C.; Powers, S.; Allen, C.M.; Gibbs
Biochemistry 32, 5167-5176, 1993
A;Title: Characterization of recombinant human farnesyl-protein transferase: cloning, ex
A;Reference number: A49274; MUID:93264431; PMID:8494894
A;Accession: A49274
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-379 <OME>
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBIP:132814)
C;Genetics:
A;Gene: GDB:FNTA
A;Cross-references: GDB:138173; OMIM:134635
A;Map position: 8p22-8q11
C;Keywords: heterodimer

Query Match 87.2%; Score 34; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
|||||

DB 270 VPHNES 275

RESULT 9
A96238
biof protein (AF311738) [imported] - Agrobacterium tumefaciens (strain C58, Cerson)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: A96238
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: A96238
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-381 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89427.1; PID:GL5159286; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L1707
A;Map position: linear chromosome
C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology

Query Match 84.6%; Score 33; DB 2; Length 381;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
|||||

DB 140 VPHNEVE 146

RESULT 10
AB3048
8-amino-7-oxononanoate synthase [imported] - Agrobacterium tumefaciens (strain C58, Dup
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AB3048
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3048
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-381 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAI44800.1; PID:GI7742441; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: bioF
A;Map position: linear chromosome
C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology

Query Match 84.6%; Score 33; DB 2; Length 381;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
|||||

DB 140 VPHNEVE 146

RESULT 11
T51269
hypothetical protein T8M16_140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C;Accession: T51269
R;Benes, V.; Wurmbach, E.; Drzonek, H.; Ansorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25346
A;Accession: T51269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <BEN>
A;Cross-references: EMBL:AL390921
A;Experimental source: cultivar Columbia; BAC clone T8M16
C;Genetics:
A;Map position: 3
A;Introns: 150/1
A;Note: T8M16_140

Query Match 82.1%; Score 32; DB 2; Length 332;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
|||||

DB 88 LPHNEGE 94

RESULT 12

S17715
transcription activator TGala - tobacco
C;Species: Nicotiana sp. (tobacco)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Dec-1998
C;Accession: S17715
R;Fromm, H.; Katagiri, F.; Chua, N.H.
Mol. Gen. Genet. 229, 181-188, 1991
A;Title: The tobacco transcription activator TGala binds to a sequence in the 5' upstream
A;Reference number: S17715; MUID:92017650; PMID:1921969
A;Accession: S17715
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-372 <PRO>
A;Title: he sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 355-Glu
C;Superfamily: fos/jun DNA-binding domain homology
F;81-124/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 82.1%; Score 32; DB 2; Length 372;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
|||||
DB 60 PHNETE 65

RESULT 13

T37923
hypothetical protein SPAC18G6.10 - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37923
R;Connor, R.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Waleh, S.V.; Wood, V.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z21754
A;Accession: T37923
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-688 <CON>
A;Cross-references: EMBL:Z68188; PIDN:CAA92388.1; GSPDB:GN00066; SPDB:SPAC18G6.10
A;Experimental source: strain 972h-; cosmid c18G6
C;Genetics:
A;Gene: SPDB:SPAC18G6.10
A;Map position: 1

Query Match 82.1%; Score 32; DB 2; Length 688;
Best Local Similarity 83.3%; Pred. No. 1.le+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNESE 7
|||||
DB 97 PHNETE 102

RESULT 14

D97132
uncharacterized phage related protein [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: D97132
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97132
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1819 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79847.1; PID:gi5024862; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824

A;Gene: CAC1883

Query Match 82.1%; Score 32; DB 2; Length 1819;
Best Local Similarity 57.1%; Pred. No. 3.le+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
|||||
DB 1289 IPHNDSQ 1295

RESULT 15

G84323
50S ribosomal protein L18p [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84323
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84323
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <STO>
A;Cross-references: GB:AE004437; NID:gi0581178; PIDN:AAG19955.1; GSPDB:GN00138
C;Genetics:
C;Gene: rpl18p
C;Superfamily: rat ribosomal protein L5

Query Match 79.5%; Score 31; DB 2; Length 183;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
|||||
DB 130 VPHNDS 135

Search completed: May 6, 2003, 15:03:31
Job time: 14.25 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 15:01:10 ; Search time 14 Seconds
(without alignments)
43.143 Million cell updates/sec

Title: US-09-851-422B-8

Perfect score: 39

Sequence: 1 VPHNESE 7

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA*
1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	357	9	US-09-809-391-755
2	34	87.2	394	10	US-09-925-302-722
3	31	79.5	316	8	US-08-945-749-4
4	31	79.5	316	8	US-08-945-749-5
5	31	79.5	648	8	US-10-081-051-22
6	30	76.9	315	8	US-08-945-749-6
7	30	76.9	419	9	US-10-182-263-1
8	30	76.9	419	9	US-10-182-263-3
9	30	76.9	419	9	US-10-182-263-4
10	30	76.9	419	9	US-10-182-263-5
11	30	76.9	419	9	US-10-182-263-6
12	30	76.9	419	9	US-09-978-317A-4
13	30	76.9	423	9	US-09-344-882-8
14	30	76.9	448	9	US-09-866-570A-56
15	30	76.9	448	10	US-09-866-572A-56
16	30	76.9	461	9	US-10-182-263-2
17	30	76.9	461	9	US-09-978-317A-2
18	30	76.9	491	9	US-10-029-180-106
19	30	76.9	610	10	US-09-764-898-201

20 74.4 29 10 US-09-864-761-48731 Sequence 48731, A
21 74.4 204 10 US-09-815-242-11657 Sequence 11657, A
22 74.4 302 9 US-09-866-570A-10 Sequence 10, Appl
23 74.4 302 10 US-09-866-572A-10 Sequence 10, Appl
24 74.4 443 9 US-09-866-570A-50 Sequence 50, Appl
25 74.4 443 10 US-09-866-572A-50 Sequence 50, Appl
26 74.4 498 9 US-09-893-519A-33 Sequence 33, Appl
27 74.4 650 10 US-09-801-368-430 Sequence 430, App
28 74.4 890 9 US-10-101-464A-958 Sequence 958, App
29 74.4 1087 10 US-09-918-909-24 Sequence 24, Appl
30 74.4 1134 9 US-09-836-392-16 Sequence 16, Appl
31 74.4 1173 9 US-10-135-322-19 Sequence 19, Appl
32 74.4 1176 9 US-09-918-508-2 Sequence 2, Appl
33 74.4 1237 10 US-09-862-027-78 Sequence 78, Appl
34 74.4 1270 9 US-10-101-464A-979 Sequence 979, Appl
35 71.8 150 10 US-09-925-301-1372 Sequence 1372, Ap
36 71.8 167 10 US-09-764-864-1138 Sequence 1138, Ap
37 71.8 197 9 US-09-738-626-5543 Sequence 5543, Ap
38 71.8 400 9 US-09-738-626-6579 Sequence 6579, Ap
39 71.8 630 9 US-10-101-464A-71 Sequence 71, Appl
40 71.8 804 10 US-09-815-242-4982 Sequence 4982, Ap
41 71.8 820 10 US-09-815-242-10771 Sequence 10771, A
42 71.8 1172 9 US-09-712-363-176 Sequence 176, App
43 71.8 1620 10 US-09-827-949-2 Sequence 2, Appl
44 69.2 20 10 US-09-839-884-58 Sequence 58, Appl
45 69.2 45 9 US-09-764-891-4164 Sequence 4164, Ap

ALIGNMENTS

RESULT 1
US-09-809-391-755
; Sequence 755, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 755
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (68)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-809-391-755

Query Match 100.0%; Score 39; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. NO. 3.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7

DB 137 VPHNESE 143

RESULT 2

US-09-925-302-722
; Sequence 722, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 722
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-722

Query Match 87.2%; Score 34; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 285 VPHNES 290

RESULT 3
US-08-945-749-4
; Sequence 4, Application US/08945749
; Patent No. US2002013880A1
; GENERAL INFORMATION:
; APPLICANT: GARDNER, Richard C
; APPLICANT: MacDIARMID, Colin W
; APPLICANT: HAY, Robert J
; APPLICANT: Auckland Uniservices Limited
; APPLICANT: New Zealand Pastoral Agriculture Research Institut
; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
; FILE REFERENCE: 08/945,749
; CURRENT APPLICATION NUMBER: US/08/945,749
; CURRENT FILING DATE: 1998-01-12
; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
; EARLIER FILING DATE: 1996-05-01
; EARLIER APPLICATION NUMBER: NZ 272039
; EARLIER FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-945-749-4

Query Match 79.5%; Score 31; DB 8; Length 316;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 227 LPHNES 232

RESULT 4
US-08-945-749-5
; Sequence 5, Application US/08945749
; Patent No. US2002013880A1
; GENERAL INFORMATION:
; APPLICANT: GARDNER, Richard C
; APPLICANT: MacDIARMID, Colin W
; APPLICANT: HAY, Robert J
; APPLICANT: Auckland Uniservices Limited
; APPLICANT: New Zealand Pastoral Agriculture Research Institut
; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
; FILE REFERENCE: 08/945,749
; CURRENT APPLICATION NUMBER: US/08/945,749
; CURRENT FILING DATE: 1998-01-12
; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
; EARLIER FILING DATE: 1996-05-01
; EARLIER APPLICATION NUMBER: NZ 272039

; EARLIER FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-08-945-749-5

Query Match 79.5%; Score 31; DB 8; Length 316;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 227 LPHNES 232

RESULT 5
US-10-081-051-22
; Sequence 22, Application US/10081051
; Publication No. US2003004422A1
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Whitmore, William M.
; APPLICANT: Kamper, Sandra M.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Ganta, Roman R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Mwangi, Duncan M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Mahan, Suman M.
; TITLE OF INVENTION: Ehrlichia ruminantium Polypeptides, Antigens, Polynucleotides, a
; FILE REFERENCE: UP-299XC1
; CURRENT APPLICATION NUMBER: US/10/081,051
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/269,944
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 22
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)
US-10-081-051-22

Query Match 79.5%; Score 31; DB 9; Length 648;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPHNESE 7
Db 329 MPHNESE 335

RESULT 6
US-08-945-749-6
; Sequence 6, Application US/08945749
; Patent No. US2002013880A1
; GENERAL INFORMATION:
; APPLICANT: GARDNER, Richard C
; APPLICANT: MacDIARMID, Colin W
; APPLICANT: HAY, Robert J
; APPLICANT: Auckland Uniservices Limited
; APPLICANT: New Zealand Pastoral Agriculture Research Institut
; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
; FILE REFERENCE: 08/945,749
; CURRENT APPLICATION NUMBER: US/08/945,749
; CURRENT FILING DATE: 1998-01-12
; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
; EARLIER FILING DATE: 1996-05-01
; EARLIER APPLICATION NUMBER: NZ 272039

; EARLIER FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Hin
US-08-945-749-6

Query Match 76.9%; Score 30; DB 8; Length 315;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PHNES 6
DB 227 PHNES 231

RESULT 7
US-10-182-263-1
; Sequence 1, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-1

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNE 5
DB 326 VPHNE 330

RESULT 8
US-10-182-263-3
; Sequence 3, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 419

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-3

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNE 5
DB 326 VPHNE 330

RESULT 9
US-10-182-263-4
; Sequence 4, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-4

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNE 5
DB 326 VPHNE 330

RESULT 10
US-10-182-263-5
; Sequence 5, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-5

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNE 5
|||||
Db 326 VPHNE 330

RESULT 11
US-10-182-263-6
; Sequence 6, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189199
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-6

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNE 5
|||||
Db 326 VPHNE 330

RESULT 12
US-09-978-917A-4
; Sequence 4, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Aps; Maxygen Holdings
; TITLE OF INVENTION: Protein C or activated protein C-like molecules
; FILE REFERENCE: G219us310 - protein C
; CURRENT APPLICATION NUMBER: US/09/978,917A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-917A-4

Query Match 76.9%; Score 30; DB 9; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNE 5
|||||
Db 326 VPHNE 330

RESULT 13
US-09-344-882-8
; Sequence 8, Application US/09344882
; Patent No. US20020162137A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurtele, Eve S

; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; CURRENT FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 8
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-8

Query Match 76.9%; Score 30; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNE 5
|||||
Db 111 VPHNE 115

RESULT 14
US-09-866-570A-56
; Sequence 56, Application US/09866570A
; Patent No. US20020168745A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,570A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/457,046B
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-570A-56

Query Match 76.9%; Score 30; DB 9; Length 448;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNES 6
|||||
Db 278 IPHNEN 283

RESULT 15
US-09-866-572A-56
; Sequence 56, Application US/09866572A
; Patent No. US20020138859A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,572A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/457,046

; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-572A-56

Query Match 76.9%; Score 30; DB 10; Length 448;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPHNES 6
:||||:
Db 278 IPHNEN 283

Search completed: May 6, 2003, 15:04:54
Job time : 15.3333 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:30 ; Search time 11.2 Seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-851-422B-8
Perfect score: 39
Sequence: 1 VPHNESE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	357	4	US-09-149-476-755
2	34	87.2	377	1	US-07-863-169A-1
3	34	87.2	377	2	US-08-429-964-1
4	34	87.2	377	3	US-07-935-087-1
5	34	87.2	377	5	PCT-US93-08062-1
6	34	87.2	379	1	US-07-863-169A-5
7	34	87.2	379	2	US-08-424-268-8
8	34	87.2	379	2	US-08-429-964-5
9	34	87.2	379	3	US-07-935-087-5
10	34	87.2	379	5	PCT-US93-08062-5
11	34	87.2	379	5	PCT-US93-10442-8
12	34	87.2	819	2	US-08-424-268-20
13	34	87.2	819	5	PCT-US93-10442-20
14	31	79.5	146	1	US-08-173-510B-88
15	31	79.5	146	1	US-08-458-218-86
16	31	79.5	146	2	US-08-450-497-88
17	30	76.9	250	4	US-08-944-483-51
18	30	76.9	261	6	5270178-5
19	30	76.9	261	6	5270178-19
20	30	76.9	261	6	5270178-20
21	30	76.9	261	6	5270178-21
22	30	76.9	262	1	US-07-720-189-1
23	30	76.9	409	4	US-09-065-872-2
24	30	76.9	409	4	US-09-667-570A-2
25	30	76.9	410	4	US-09-065-872-1
26	30	76.9	410	4	US-09-667-570A-1
27	30	76.9	419	1	US-08-295-411-1

28	30	76.9	419	2	US-08-955-471-1
29	30	76.9	419	4	US-09-667-570A-3
30	30	76.9	419	5	PCT-US92-10242-1
31	30	76.9	448	4	US-09-457-046B-56
32	30	76.9	460	2	US-08-756-506-2
33	30	76.9	460	2	US-08-756-506-4
34	30	76.9	460	6	5270178-13
35	30	76.9	460	6	5270178-14
36	30	76.9	460	6	5270178-15
37	30	76.9	460	6	5270178-16
38	30	76.9	461	6	5225537-2
39	30	76.9	461	6	5270178-2
40	30	76.9	461	6	5270178-17
41	30	76.9	461	6	5270178-18
42	30	76.9	758	1	US-08-258-188-2
43	30	76.9	758	1	US-08-526-813-2
44	30	76.9	758	5	PCT-US95-08554-2
45	30	76.9	1349	3	US-08-938-291A-6

ALIGNMENTS

RESULT 1

US-09-149-476-755
; Sequence 755, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002PI
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23

Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 56, Appli
Sequence 2, Appli
Sequence 4, Appli
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Patent No. 5270178
Sequence 2, Appli
Sequence 2, Appli
Sequence 6, Appli

1	EARLIER APPLICATION NUMBER: 60/047,500
2	EARLIER FILING DATE: 1997-05-23
3	EARLIER APPLICATION NUMBER: 60/047,587
4	EARLIER FILING DATE: 1997-05-23
5	EARLIER APPLICATION NUMBER: 60/047,492
6	EARLIER FILING DATE: 1997-05-23
7	EARLIER APPLICATION NUMBER: 60/047,598
8	EARLIER FILING DATE: 1997-05-23
9	EARLIER APPLICATION NUMBER: 60/047,613
10	EARLIER FILING DATE: 1997-05-23
11	EARLIER APPLICATION NUMBER: 60/047,582
12	EARLIER FILING DATE: 1997-05-23
13	EARLIER APPLICATION NUMBER: 60/047,596
14	EARLIER FILING DATE: 1997-05-23
15	EARLIER APPLICATION NUMBER: 60/047,612
16	EARLIER FILING DATE: 1997-05-23
17	EARLIER APPLICATION NUMBER: 60/047,632
18	EARLIER FILING DATE: 1997-05-23
19	EARLIER APPLICATION NUMBER: 60/047,601
20	EARLIER FILING DATE: 1997-05-23
21	EARLIER APPLICATION NUMBER: 60/043,580
22	EARLIER FILING DATE: 1997-04-11
23	EARLIER APPLICATION NUMBER: 60/043,568
24	EARLIER FILING DATE: 1997-04-11
25	EARLIER APPLICATION NUMBER: 60/043,314
26	EARLIER FILING DATE: 1997-04-11
27	EARLIER APPLICATION NUMBER: 60/043,569
28	EARLIER FILING DATE: 1997-04-11
29	EARLIER APPLICATION NUMBER: 60/043,311
30	EARLIER FILING DATE: 1997-04-11
31	EARLIER APPLICATION NUMBER: 60/043,671
32	EARLIER FILING DATE: 1997-04-11
33	EARLIER APPLICATION NUMBER: 60/043,674
34	EARLIER FILING DATE: 1997-04-11
35	EARLIER APPLICATION NUMBER: 60/043,669
36	EARLIER FILING DATE: 1997-04-11
37	EARLIER APPLICATION NUMBER: 60/043,312
38	EARLIER FILING DATE: 1997-04-11
39	EARLIER APPLICATION NUMBER: 60/043,313
40	EARLIER FILING DATE: 1997-04-11
41	EARLIER APPLICATION NUMBER: 60/043,672
42	EARLIER FILING DATE: 1997-04-11
43	EARLIER APPLICATION NUMBER: 60/043,315
44	EARLIER FILING DATE: 1997-04-11
45	EARLIER APPLICATION NUMBER: 60/048,974
46	EARLIER FILING DATE: 1997-06-06
47	EARLIER APPLICATION NUMBER: 60/056,886
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,877
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,889
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,893
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,630
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,878
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,662
60	EARLIER FILING DATE: 1997-08-22
61	EARLIER APPLICATION NUMBER: 60/056,872
62	EARLIER FILING DATE: 1997-08-22
63	EARLIER APPLICATION NUMBER: 60/056,882
64	EARLIER FILING DATE: 1997-08-22
65	EARLIER APPLICATION NUMBER: 60/056,637
66	EARLIER FILING DATE: 1997-08-22
67	EARLIER APPLICATION NUMBER: 60/056,903
68	EARLIER FILING DATE: 1997-08-22
69	EARLIER APPLICATION NUMBER: 60/056,888
70	EARLIER FILING DATE: 1997-08-22
71	EARLIER APPLICATION NUMBER: 60/056,879
72	EARLIER FILING DATE: 1997-08-22
73	EARLIER APPLICATION NUMBER: 60/056,880

1	EARLIER FILING DATE: 1997-08-22
2	EARLIER APPLICATION NUMBER: 60/056,894
3	EARLIER FILING DATE: 1997-08-22
4	EARLIER APPLICATION NUMBER: 60/056,911
5	EARLIER FILING DATE: 1997-08-22
6	EARLIER APPLICATION NUMBER: 60/056,936
7	EARLIER FILING DATE: 1997-08-22
8	EARLIER APPLICATION NUMBER: 60/056,874
9	EARLIER FILING DATE: 1997-08-22
10	EARLIER APPLICATION NUMBER: 60/056,910
11	EARLIER FILING DATE: 1997-08-22
12	EARLIER APPLICATION NUMBER: 60/056,864
13	EARLIER FILING DATE: 1997-08-22
14	EARLIER APPLICATION NUMBER: 60/056,631
15	EARLIER FILING DATE: 1997-08-22
16	EARLIER APPLICATION NUMBER: 60/056,845
17	EARLIER FILING DATE: 1997-08-22
18	EARLIER APPLICATION NUMBER: 60/056,892
19	EARLIER FILING DATE: 1997-08-22
20	EARLIER APPLICATION NUMBER: 60/057,761
21	EARLIER FILING DATE: 1997-08-22
22	EARLIER APPLICATION NUMBER: 60/047,595
23	EARLIER FILING DATE: 1997-05-23
24	EARLIER APPLICATION NUMBER: 60/047,599
25	EARLIER FILING DATE: 1997-05-23
26	EARLIER APPLICATION NUMBER: 60/047,588
27	EARLIER FILING DATE: 1997-05-23
28	EARLIER APPLICATION NUMBER: 60/047,585
29	EARLIER FILING DATE: 1997-05-23
30	EARLIER APPLICATION NUMBER: 60/047,586
31	EARLIER FILING DATE: 1997-05-23
32	EARLIER APPLICATION NUMBER: 60/047,590
33	EARLIER FILING DATE: 1997-05-23
34	EARLIER APPLICATION NUMBER: 60/047,594
35	EARLIER FILING DATE: 1997-05-23
36	EARLIER APPLICATION NUMBER: 60/047,589
37	EARLIER FILING DATE: 1997-05-23
38	EARLIER APPLICATION NUMBER: 60/047,593
39	EARLIER FILING DATE: 1997-05-23
40	EARLIER APPLICATION NUMBER: 60/047,614
41	EARLIER FILING DATE: 1997-05-23
42	EARLIER APPLICATION NUMBER: 60/043,578
43	EARLIER FILING DATE: 1997-04-11
44	EARLIER APPLICATION NUMBER: 60/043,576
45	EARLIER FILING DATE: 1997-04-11
46	EARLIER APPLICATION NUMBER: 60/047,501
47	EARLIER FILING DATE: 1997-05-23
48	EARLIER APPLICATION NUMBER: 60/043,670
49	EARLIER FILING DATE: 1997-04-11
50	EARLIER APPLICATION NUMBER: 60/056,632
51	EARLIER FILING DATE: 1997-08-22
52	EARLIER APPLICATION NUMBER: 60/056,664
53	EARLIER FILING DATE: 1997-08-22
54	EARLIER APPLICATION NUMBER: 60/056,876
55	EARLIER FILING DATE: 1997-08-22
56	EARLIER APPLICATION NUMBER: 60/056,881
57	EARLIER FILING DATE: 1997-08-22
58	EARLIER APPLICATION NUMBER: 60/056,909
59	EARLIER FILING DATE: 1997-08-22
60	EARLIER APPLICATION NUMBER: 60/056,875
61	EARLIER FILING DATE: 1997-08-22
62	EARLIER APPLICATION NUMBER: 60/056,862
63	EARLIER FILING DATE: 1997-08-22
64	EARLIER APPLICATION NUMBER: 60/056,887
65	EARLIER FILING DATE: 1997-08-22
66	EARLIER APPLICATION NUMBER: 60/056,908
67	EARLIER FILING DATE: 1997-08-22
68	EARLIER APPLICATION NUMBER: 60/048,964
69	EARLIER FILING DATE: 1997-06-06
70	EARLIER APPLICATION NUMBER: 60/057,650
71	EARLIER FILING DATE: 1997-09-05
72	EARLIER APPLICATION NUMBER: 60/056,884
73	EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 100.0%; Score 39; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
DB 137 VPHNESE 143

RESULT 2
US-07-863-169A-1
Sequence 1, Application US/07863169A
Patent No. 5420245
GENERAL INFORMATION:
APPLICANT: Brown, Michael S.
APPLICANT: Goldstein, Joseph L.
APPLICANT: Reiss, Yuval
TITLE OF INVENTION: Tetrapeptide-Based Inhibitors of Farnesyl
TITLE OF INVENTION: Transferase
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/863,169A
FILING DATE: 03-APR-1992

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/822,011
FILING DATE: 19-JAN-1992
CLASSIFICATION: 530
APPLICATION NUMBER: US 07/937,893
FILING DATE: 18-APR-1991
CLASSIFICATION: 530
APPLICATION NUMBER: US 615,715
FILING DATE: 20-NOV-1990
CLASSIFICATION: 530
APPLICATION NUMBER: US 510,706
FILING DATE: 18-APR-1990
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: US/07/297/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-07-863-169A-1
Query Match 87.2%; Score 34; DB 1; Length 377;

Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
DB 270 VPHNES 275

RESULT 3
US-08-429-964-1
Sequence 1, Application US/08429964
Patent No. 5962243
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
APPLICANT: REISS, YUVAL
APPLICANT: JAMES, GUY L.
TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
TITLE OF INVENTION: TRANSFERASE INHIBITORS
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,964
FILING DATE: 27-APR-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,625
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/822,011
FILING DATE: ABANDONED
CLASSIFICATION: 435
APPLICATION NUMBER: PCT/US/91/02650
FILING DATE: 18-APR-1991
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/615,715
FILING DATE: 20-NOV-1990
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/510,706
FILING DATE: 18-APR-1990 (ABANDONED)

ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: US/08/432/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-429-964-1
Query Match 87.2%; Score 34; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

Db 270 VPHNES 275
|||||

RESULT 4

US-07-935-087-1

; Sequence 1, Application US/07935087

; Patent No. 6083917

; GENERAL INFORMATION:

; APPLICANT: BROWN, MICHAEL S.

; APPLICANT: GOLDSTEIN, JOSEPH L.

; APPLICANT: REISS, YUVAL

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR

; TITLE OF INVENTION: THE IDENTIFICATION,

; TITLE OF INVENTION: CHARACTERIZATION,

; TITLE OF INVENTION: AND INHIBITION OF FARNESYL

; TITLE OF INVENTION: PROTEIN TRANSFERASE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARNOLD, WHITE & DURKEE

; STREET: P.O. BOX 4433

; CITY: HOUSTON

; STATE: TEXAS

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,087

; FILING DATE: 19920824

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/822,011

; FILING DATE: 01/16/92

; ATTORNEY/AGENT INFORMATION:

; NAME: PARKER, DAVID L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: UTSD:269/PAR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512-320-7200

; TELEFAX: 512-474-7577

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 377 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-935-087-1

Query Match

Best Local Similarity 87.2%; Score 34; DB 3; Length 377;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

Db 270 VPHNES 275

RESULT 5

PCT-US93-08062-1

; Sequence 1, Application PC/TUS9308062

; GENERAL INFORMATION:

; APPLICANT:

; SEQUENCE CHARACTERISTICS:

; SEQUENCE CHARACTERISTICS: BROWN, MICHAEL S.

; SEQUENCE CHARACTERISTICS: GOLDSTEIN, JOSEPH L.

; SEQUENCE CHARACTERISTICS: REISS, YUVAL

; SEQUENCE CHARACTERISTICS: MARSTERS, JR., JAMES C.

; ADDRESSEE: METHODS AND COMPOSITIONS FOR

; ADDRESSEE: THE IDENTIFICATION,

; ADDRESSEE: CHARACTERIZATION AND

; ADDRESSEE: INHIBITION OF

; ADDRESSEE: FARNESYLTRANSFERASE

; NUMBER OF SEQUENCES: 71

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARNOLD, WHITE & DURKEE

; STREET: P.O. BOX 4433

; CITY: HOUSTON

; STATE: TEXAS

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK/ASKII

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/08062

; FILING DATE: AUGUST 24, 1993

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/935,087

; FILING DATE: 24 AUGUST 1992 (24.08.92)

; NAME: UNKNOWN

; ATTORNEY/AGENT INFORMATION:

; NAME: PARKER, DAVID L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: UTFD377PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512-320-7200

; TELEFAX: 512-474-7577

; TELEX: NOT APPLICABLE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 377 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

PCT-US93-08062-1

Query Match

Best Local Similarity 87.2%; Score 34; DB 5; Length 377;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

Db 270 VPHNES 275

RESULT 6

US-07-863-169A-5

; Sequence 5, Application US/07863169A

; Patent No. 5420245

; GENERAL INFORMATION:

; APPLICANT: Brown, Michael S.

; APPLICANT: Goldstein, Joseph L.

; TITLE OF INVENTION: Tetrapeptide-Based Inhibitors of Farnesyl

; TITLE OF INVENTION: Transferase

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: United States of America

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US 07/863,169A
;; FILING DATE: 03-APR-1992
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/822,011
;; FILING DATE: 19-JAN-1992
;; CLASSIFICATION: 530
;; APPLICATION NUMBER: US 07/937,893
;; FILING DATE: 18-APR-1991
;; CLASSIFICATION: 530
;; APPLICATION NUMBER: US 615,715
;; FILING DATE: 20-NOV-1990
;; CLASSIFICATION: 530
;; APPLICATION NUMBER: US 510,706
;; FILING DATE: 18-APR-1990
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, David L.
;; REGISTRATION NUMBER: 32,165
;; REFERENCE/DOCKET NUMBER: UTSD:237/PAR
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (512) 418-3000
;; TELEFAX: (713) 789-2679
;; TELEX: 79-0924
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 379 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-07-863-169A-5

Query Match 87.2%; Score 34; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPHNES 6
|||
DB 270 VPHNES 275

RESULT 7
US-08-424-268-8
; Sequence 8, Application US/08424268
; Patent No. 5821118
; GENERAL INFORMATION:
; APPLICANT: Omer, Charles A
; APPLICANT: Diehl, Ronald E
; APPLICANT: Gibbs, Jackson B
; APPLICANT: Kohl, Nancy E
; TITLE OF INVENTION: Assay for Inhibitors of Farnesyl-Protein
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O.Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Power Mac
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,268
; FILING DATE: 4/24/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muchard, David A
; REGISTRATION NUMBER: 35,297

;; REFERENCE/DOCKET NUMBER: 18858PC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 594-3903
;; TELEFAX: (908) 594-4720
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 379-amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; FRAGMENT TYPE: N-terminal
;; US-08-424-268-8

Query Match 87.2%; Score 34; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPHNES 6
|||
DB 270 VPHNES 275

RESULT 8
US-08-429-964-5
; Sequence 5, Application US/08429964
; Patent No. 5962243
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; APPLICANT: JAMES, GUY L.
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
; TITLE OF INVENTION: TRANSFERASE INHIBITORS
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,964
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,625
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/822,011
; FILING DATE: ABANDONED
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/US/91/02650
; FILING DATE: 18-APR-1991
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/615,715
; FILING DATE: 20-NOV-1990
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/510,706
; FILING DATE: 18-APR-1990 (ABANDONED)
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:432/PAR
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-429-964-5

Query Match 87.2%; Score 34; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 270 VPHNES 275

RESULT 9
US-07-935-087-5
Sequence 5, Application US/07935087
Patent No. 6083917
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
APPLICANT: REISS, YUVAL
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: THE IDENTIFICATION,
TITLE OF INVENTION: CHARACTERIZATION,
TITLE OF INVENTION: AND INHIBITION OF FARNESYL
TITLE OF INVENTION: PROTEIN TRANSFERASE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/935,087
FILING DATE: 19920824
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/822,011
FILING DATE: 01/16/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:269/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-07-935-087-5

Query Match 87.2%; Score 34; DB 3; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 270 VPHNES 275

RESULT 10
PCT-US93-08062-5
Sequence 5, Application PC/TUS9308062
GENERAL INFORMATION:
APPLICANT:
SEQUENCE CHARACTERISTICS:
SEQUENCE CHARACTERISTICS: BROWN, MICHAEL S.
SEQUENCE CHARACTERISTICS: GOLDSTEIN, JOSEPH L.
SEQUENCE CHARACTERISTICS: REISS, YUVAL
SEQUENCE CHARACTERISTICS: MARSTERS, JR., JAMES C.
ADDRESSEE: METHODS AND COMPOSITIONS FOR
ADDRESSEE: THE IDENTIFICATION,
ADDRESSEE: CHARACTERIZATION AND
ADDRESSEE: INHIBITION OF
ADDRESSEE: FARNESYLTRANSFERASE
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK/ASKII
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08062
FILING DATE: AUGUST 24, 1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/935,087
FILING DATE: 24 AUGUST 1992 (24.08.92)
NAME: UNKNOWN
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD377PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-08062-5

Query Match 87.2%; Score 34; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 270 VPHNES 275

RESULT 11
PCT-US93-10442-8
Sequence 8, Application PC/TUS9310442
GENERAL INFORMATION:
APPLICANT: Omer,
APPLICANT: Diehl, Ronald E
APPLICANT: Gibbs, Jackson B

APPLICANT: Kohl, Nancy E
TITLE OF INVENTION: Assay for Inhibitors of Farnesyl-Protein
TRANSFERASE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10442
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/968,782
FILING DATE: 10/30/92
ATTORNEY/AGENT INFORMATION:
NAME: Muthard, David A
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 18858
TELEPHONE: (908) 594-3903
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
PCT-US93-10442-8

Query Match 87.2%; Score 34; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPINES 6

Db 270 VPINES 275

RESULT 12
US-08-424-268-20
Sequence 20, Application US/08424268
Patent No. 5821118
GENERAL INFORMATION:
APPLICANT: Omer, Charles A
APPLICANT: Diehl, Ronald E
APPLICANT: Gibbs, Jackson B
APPLICANT: Kohl, Nancy E
TITLE OF INVENTION: Assay for Inhibitors of Farnesyl-Protein
TRANSFERASE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Mac
OPERATING SYSTEM: System 7.5.3
SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,268
FILING DATE: 4/24/95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muthard, David A
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 18858PC
TELEPHONE: (908) 594-3903
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 819 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
US-08-424-268-20
Query Match 87.2%; Score 34; DB 2; Length 819;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPINES 6
Db 710 VPINES 715

RESULT 13
PCT-US93-10442-20
Sequence 20, Application PC/TUS9310442
GENERAL INFORMATION:
APPLICANT: Omer, Charles A
APPLICANT: Diehl, Ronald E
APPLICANT: Gibbs, Jackson B
APPLICANT: Kohl, Nancy E
TITLE OF INVENTION: Assay for Inhibitors of Farnesyl-Protein
TRANSFERASE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10442
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/968,782
FILING DATE: 10/30/92
ATTORNEY/AGENT INFORMATION:
NAME: Muthard, David A
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 18858
TELEPHONE: (908) 594-3903
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 819 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
PCT-US93-10442-20

Query Match 87.2%; Score 34; DB 5; Length 819;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPINES 6
DB 710 VPINES 715

RESULT 14
US-08-173-510B-88
Sequence 88, Application US/08173510B
Patent No. 5747296
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE, ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,510B
FILING DATE: 23-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOV-1993
APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993
APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992
FILING DATE: 11-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 205/073

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE

US-08-173-510B-88

Query Match 79.5%; Score 31; DB 1; Length 146;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PHNESE 7
DB 25 PHNEGE 30

RESULT 15
US-08-458-218-86
Sequence 86, Application US/08458218
Patent No. 5789178
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,218
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOVEMBER-1993

APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993

APPLICATION NUMBER: 07/881,721
FILING DATE: 11-MAY-1992

APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992

ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 203/226
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440

TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:

LENGTH: 146 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE

US-08-458-218-86

Query Match 79.5%; Score 31; DB 1; Length 146;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PHNESE 7
DB 25 PHNEGE 30

Search completed: May 6, 2003, 15:04:07
Job time: 12.2 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:44 ; Search time 31.85 Seconds
(without alignments)
29.286 Million cell updates/sec

Title: US-09-851-422B-8

Perfect score: 39
Sequence: 1 VPHNESE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	444	AAU30756	Novel human secret
2	36	92.3	242	AAU30756	Arabidopsis thalia
3	36	92.3	245	AAU30756	Arabidopsis thalia
4	36	92.3	315	AAU30756	Arabidopsis thalia
5	34	87.2	13	AA194717	Farnesyl-protein t
6	34	87.2	293	AA194717	Arabidopsis thalia
7	34	87.2	377	AA194717	Farnesylproteins
8	34	87.2	377	AA194717	Rat farnesyl prote
9	34	87.2	377	AA194717	Rat farnesyl trans
10	34	87.2	377	AA194717	Mouse ischaemic co

11	34	87.2	379	15	AA49734	Farnesyltransferas
12	34	87.2	379	15	AA49734	Alpha subunit of h
13	34	87.2	379	16	AA49734	Human farnesyl pro
14	34	87.2	379	17	AAU04431	Human farnesyl tra
15	34	87.2	379	21	AB008436	Protein sequence 2
16	34	87.2	380	21	AAU07150	Human geranylgeran
17	34	87.2	394	21	AA058384	Lung cancer associ
18	34	87.2	394	21	AA058384	Drosophila melanog
19	32	82.1	198	22	ABG21522	Novel human diagno
20	32	82.1	332	21	AA055461	Arabidopsis thalia
21	32	82.1	341	21	AA055461	Arabidopsis thalia
22	32	82.1	1639	22	AB059281	Drosophila melanog
23	31	79.5	71	23	ABP03959	Human ORFX protein
24	31	79.5	87	18	AAW55657	H. pylori ORF 11ae
25	31	79.5	100	18	AAW55657	H. pylori ORF 06cp
26	31	79.5	146	14	AA052989	Canine hookworm Ne
27	31	79.5	146	20	AAV23597	Canine hookworm ne
28	31	79.5	534	22	AB063601	Drosophila melanog
29	30	76.9	9	19	AAW66095	Protein C peptide
30	30	76.9	78	12	AA015744	R18 antigen of HTL
31	30	76.9	233	23	AB083175	Streptococcus suis
32	30	76.9	235	22	ABG02914	Novel human diagno
33	30	76.9	258	12	AA013048	Protein C heavy ch
34	30	76.9	261	14	AA030723	Protein C heavy ch
35	30	76.9	262	12	AA013047	Protein C heavy ch
36	30	76.9	262	12	AA013047	Protein C heavy ch
37	30	76.9	262	12	AA013047	Human protein C ca
38	30	76.9	262	12	AA013047	Human protein C ca
39	30	76.9	262	12	AA013047	Human protein C ca
40	30	76.9	262	12	AA013047	Human protein C ca
41	30	76.9	262	12	AA013047	Human protein C ca
42	30	76.9	262	12	AA013047	Human protein C ca
43	30	76.9	263	15	AA062654	Activated modified
44	30	76.9	272	21	AA021970	Arabidopsis thalia
45	30	76.9	329	21	AA021970	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAU30756
ID AAU30756 standard; Protein; 444 AA.
XX
AC AAU30756;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #1247.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
(HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy

XX Claim 20; Page 340; 765pp; English.

PS The invention relates to novel human secreted polypeptides. The

XX polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising

CC the nucleic acids encoding the polypeptides and cells genetically

CC engineered to express them are also useful for producing the proteins.

CC The proteins are useful in genetic vaccination, testing and

CC therapy, and can be used as nutritional supplements. They may be used to

CC increase stem cell proliferation; to regulate haematopoiesis; and in

CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;

CC immune suppression and/or stimulation; as anti-inflammatory agents; and

CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid

XX sequences of novel human secreted proteins of the invention.

SQ Sequence 444 AA;

Query Match 100.0%; Score 39; DB 22; Length 444;

Best Local Similarity 100.0%; Pred. No. 6.7;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7

Db 209 VPHNESE 215

|||||

RESULT 2

AAU09225

ID AAG09225 standard; Protein; 242 AA.

XX AC AAG09225;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7073.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132048.

PR 05-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139763.

PR 22-JUN-1999; 99US-0139817.

PR 23-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144684.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147933.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 92.3%; Score 36; DB 21; Length 242;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
|||||:
Db 145 VPHNESD 151

RESULT 3

AAG09224
ID AAG09224 standard; Protein; 245 AA.

XX AAG09224;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 7072.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143354.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 92.3%; Score 36; DB 21; Length 245;

Best Local Similarity 85.7%; Pred. No. 15;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7

Db 148 VPHNESD 154

RESULT 4

AAG09223
 ID AAG09223 standard; Protein; 315 AA.

XX AC AAG09223;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7071.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.

```
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.
Query Match 92.3%; Score 36; DB 21; Length 315;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNESE 7
Db 218 VPHNESD 224
|||||:

RESULT 5
AAR14717
ID AAR14717 standard; Protein; 13 AA.
XX
AC AAR14717;
XX
DT 29-JAN-1992 (first entry)
XX
DE Farnesyl-protein transferase alpha subunit (6).
XX
KW Farnesyl; transferase; FT; inhibitor; p21ras; rat; cancer.
XX
OS Rattus rattus.
XX
PN WO9116340-A.
XX
PD 31-OCT-1991.
XX
PF 18-APR-1991; 91WO-US02650.
XX
PR 20-NOV-1990; 90US-0615715.
PR 18-APR-1990; 90US-0510706.
XX
PA (TEXA ) UNIV OF TEXAS SYST.
XX
PI Brown MS, Goldstein JL, Reiss Y;
XX
WPI; 1991-339750/46.
Compen. comprising purified farnesyl-protein transferase - used
to inhibit attachment of farnesyl moiety to RAS protein in
malignant cells and to treat cancer
XX
PS Disclosure; Page 8; 87pp; English.
XX
CC FT comprises two subunits each having a mol. wt. of 45-50 kD (SDS-PAGE).
The tryptic digests of the alpha unit are represented in AAR14712-17,
and of the beta unit in AAR14718-22.
XX
CC The FT enzyme can be used to isolate inhibitors. Such inhibitors
can be used to block the attachment of farnesyl gps. to ras
proteins in malignant cells of patients suffering with cancer or
pre-cancerous states to treat or palliate the cancer.
XX
CC See also AAR14711-23, AAR15751-81 and AAQ14541-47.
SQ Sequence 13 AA;

Query Match 87.2%; Score 34; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
Db 2 VPHNES 7
|||||

RESULT 6
AAG30075
ID AAG30075 standard; Protein; 293 AA.
XX
AC AAG30075;
XX
DT 17-OCT-2000 (first entry)
```


PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159838.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161320.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 87.2%; Score 34; DB 21; Length 293;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
 |||||
 Db 74 VPHNES 79

RESULT 7
 AAR49739
 ID AAR49739 standard; Protein; 377 AA.
 XX
 AC AAR49739;
 XX

DT 08-AUG-1994 (first entry)
 XX
 DE Farnesyltransferase alpha-subunit.
 XX
 KW Farnesyltransferase; FT; alpha-subunit; p21ras; ras protein;
 cancer therapy.
 XX
 OS Rattus sp.

XX
 PN WO9404561-A.
 XX
 PD 03-MAR-1994.
 XX
 PF 24-AUG-1993; 93WO-US08062.
 XX
 PR 24-AUG-1992; 92US-0935087.
 XX
 PA (GETH) GENENTECH INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Brown MS, Goldstein JL, Marsters JC, Reiss Y;
 XX
 WPI, 1994-083105/10.
 DR N-PSDB; AAQ44397.
 XX
 PT New farnesyl-transferase inhibitors - used for inhibiting
 PT attachment of a farnesyl moiety to a p21ras protein in malignant
 PT cells
 XX
 PS Disclosure; Page 116-118; 183pp; English.
 XX
 CC The cDNA (AAQ44397) and amino acid (AAR49739) sequences of rat
 CC farnesyltransferase (FT) alpha-subunit were determined. The cDNA
 CC may be used to produce recombinant FT, useful for screening
 CC potential anticancer agents that prevent expression of p21ras.
 XX
 SQ Sequence 377 AA;

Query Match 87.2%; Score 34; DB 15; Length 377;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
 |||||
 Db 270 VPHNES 275

RESULT 8
 AAR77839
 ID AAR77839 standard; Protein; 377 AA..
 XX
 AC AAR77839;
 XX
 DT 23-JAN-1996 (first entry)
 XX
 DE Rat farnesyl protein transferase alpha subunit.
 XX
 KW Farnesyl transferase; inhibitor; cancer; ras; p21.
 XX
 OS Rattus sp.
 XX
 PN US5420245-A.
 XX
 PD 30-MAY-1995.
 XX
 PF 18-APR-1990; 90US-0510706.
 XX
 PR 16-JAN-1992; 92US-0822011.
 PR 18-APR-1990; 90US-0510706.
 PR 20-NOV-1990; 90US-0615715.
 PR 03-APR-1992; 92US-0863169.
 XX
 PA (TEXA) UNIV TEXAS.
 XX
 PI Brown MS, Goldstein JL, Reiss Y;
 XX
 WPI, 1995-206308/27.
 DR N-PSDB; AAQ94410.
 XX
 PT New farnesyl transferase inhibitor peptide(s) - based on farnesyl
 PT acceptor substrate carboxy terminal sequences, used for the

PT treatment of cancer

PS Example 3; Column 47-50; 55pp; English.

XX AAR77839 is the alpha subunit of rat farnesyl transferase which
CC is involved in the farnesylation of various cellular proteins
CC including the cancer related ras proteins. It is used to produce the
CC complete farnesyl transferase molecule which is used to demonstrate
CC the effectiveness of peptide inhibitors capable of inhibiting
CC farnesyl transferases. The peptide inhibitors are useful for
CC treating cancers and ras-related cancers in particular.

XX Sequence 377 AA;

Query Match 87.2%; Score 34; DB 16; Length 377;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

DB 270 VPHNES 275

RESULT 9

AAW04427
ID AAW04427 standard; Protein; 377 AA.

XX AAW04427;

DT 30-JUL-1997 (first entry)

DE Rat farnesyl transferase enzyme alpha subunit.

XX Farnesyl transferase; inhibitor; cancer; tumour; neoplasia; prenyl;
KW ras protein; K-ras B; malignant; detection; identification.

OS Rattus rattus.

XX WO9634113-A2.

PD 31-OCT-1996.

PF 29-APR-1996; 96WO-US05969.

PR 27-APR-1995; 95US-0429964.

XX (TEXA) UNIV TEXAS SYSTEM.

PI Brown MS, Goldstein JL, James GL;

DR N-PSDB; AAT38708.

XX Assay for farnesyl transferase activity - by determining ability to
PT transfer farnesyl moiety to K-Ras B protein, partic. useful for
PT identifying inhibitors

PS Example 3; Page 137-140; 257pp; English.

XX AAW04427 shows the alpha subunit of a farnesyl transferase (Ft) enzyme
CC derived from rat brain tissue. The enzyme was used in a method
CC for identifying Ft inhibitors. The method involved screening candidate
CC compounds for the ability to inhibit the transfer of a farnesyl moiety
CC to a K-ras B protein. Ft inhibitors act by blocking the attachment of
CC prenyl groups to ras proteins in malignant cells of patients suffering
CC from cancer or precancerous states, and as such are used to treat such
CC conditions.

XX Sequence 377 AA;

Query Match 87.2%; Score 34; DB 17; Length 377;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

DB 270 VPHNES 275

RESULT 10

ABB57154
ID ABB57154 standard; Protein; 377 AA.

XX ABB57154;

DT 07-MAR-2002 (first entry)

XX Mouse ischaemic condition related protein sequence SEQ ID NO:375.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.

OS Mus musculus.

XX WO200188189-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

DR WPI; 2002-034733/04.

DR N-PSDB; ABI99441.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -

PS Claim 2; Page 1003-1004; 2690pp; English.

XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC the expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX Sequence 377 AA;

Query Match 87.2%; Score 34; DB 23; Length 377;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

DB 270 VPHNES 275

RESULT 11

AAR49734
ID AAR49734 standard; Protein; 379 AA.

XX

```

AC AAR49734;
XX
XX DT 08-AUG-1994 (first entry)
XX
XX DE Farnesyltransferase alpha-subunit.
XX
XX KW Farnesyltransferase; FT; alpha-subunit; p21ras; ras protein;
XX cancer therapy.
XX
XX OS Homo sapiens.
XX
XX PN WO9404561-A.
XX
XX PD 03-MAR-1994.
XX
XX PF 24-AUG-1993; 93WO-US08062.
XX
XX PR 24-AUG-1992; 92US-0935087.
XX
XX PA (GETH ) GENENTECH INC.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX PI Brown MS, Goldstein JL, Marsters JC, Reiss Y;
XX
XX DR WPI; 1994-083105/10.
XX N-PSDB; AAQ44395.
XX
XX PT New farnesyl-transferase inhibitors - used for inhibiting
XX attachment of a farnesyl moiety to a p21ras protein in malignant
XX cells
XX
XX PS Disclosure; Fig 26; 183pp; English.
XX
XX CC The cDNA (AAQ44395) and amino acid (AAR49734) sequences of human
XX farnesyltransferase (FT) alpha-subunit were determined. The cDNA
XX may be used to produce recombinant FT, useful for screening
XX potential anticancer agents that prevent expression of p21ras.
XX
XX SQ Sequence 379 AA;

Query Match 87.2%; Score 34; DB 15; Length 379;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
DB 270 VPHNES 275

RESULT 12
AAR54830
ID AAR54830 standard; Protein; 379 AA.
XX
XX AC AAR54830;
XX
XX DT 07-NOV-1994 (first entry)
XX
XX DE Alpha subunit of human FTFase.
XX
XX KW Farnesyl protein transferase; inhibition; farnesylation.
XX
XX OS Homo sapiens.
XX
XX PN WO9410184-A.
XX
XX PD 11-MAY-1994.
XX
XX PF 29-OCT-1993; 93WO-US10442.
XX
XX PR 30-OCT-1992; 92US-0968782.
XX
XX PA (MERI ) MERCK & CO INC.
XX

Query Match 87.2%; Score 34; DB 15; Length 379;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
DB 270 VPHNES 275

RESULT 13
AAR77841
ID AAR77841 standard; Protein; 379 AA.
XX
XX AC AAR77841;
XX
XX DT 23-JAN-1996 (first entry)
XX
XX DE Human farnesyl protein transferase alpha subunit.
XX
XX KW Farnesyl transferase; inhibitor; cancer; ras; p21.
XX
XX OS Homo sapiens.
XX
XX PN US5420245-A.
XX
XX PD 30-MAY-1995.
XX
XX PF 18-APR-1990; 90US-0510706.
XX
XX PR 16-JAN-1992; 92US-0822011.
XX
XX PR 18-APR-1990; 90US-0510706.
XX
XX PR 20-NOV-1990; 90US-0615715.
XX
XX PR 03-APR-1992; 92US-0863169.
XX
XX PA (TEXA ) UNIV TEXAS.
XX
XX PI Brown MS, Goldstein JL, Reiss Y;
XX
XX DR WPI; 1995-206308/27.
XX N-PSDB; AAQ94412.
XX
XX PT New farnesyl transferase inhibitor peptide(s) - based on farnesyl
XX acceptor substrate carboxy terminal sequences, used for the
XX treatment of cancer
XX
XX PS Example 4; Column 55-58; 55pp; English.
XX
XX CC AAR77841 is the alpha subunit of human farnesyl transferase which
XX is involved in the farnesylation of various cellular proteins
XX including the cancer related ras proteins. It is used to produce the
XX complete farnesyl transferase molecule which is used to demonstrate
XX the effectiveness of peptide inhibitors capable of inhibiting
XX farnesyl transferases. The peptide inhibitors are useful for

```

```

PI Diehl RE, Gibbs JB, Kohl NE, Omer CA;
XX
XX DR WPI; 1994-167373/20.
XX N-PSDB; AAQ64887.
XX
XX PT Mammalian farnesyl protein transferase prodn. - used in assays
XX for cpds. with inhibitory activity for the identification of
XX anticancer agents
XX
XX PS Disclosure; Fig 2; 69pp; English.
XX
XX CC The cDNA encoding the human alpha subunit of FTFase was isolated
XX from a human placental cDNA library in lambda gt11 using a bovine
XX FTFase cDNA probe. The FTFase can be used to assess the inhibitory
XX activity of a cpd. in the farnesylation of a protein substrate.
XX The assay can be used to identify anticancer agents.
XX See also AAR54829-32.
XX
XX SQ Sequence 379 AA;

Query Match 87.2%; Score 34; DB 15; Length 379;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
DB 270 VPHNES 275

RESULT 13
AAR77841
ID AAR77841 standard; Protein; 379 AA.
XX
XX AC AAR77841;
XX
XX DT 23-JAN-1996 (first entry)
XX
XX DE Human farnesyl protein transferase alpha subunit.
XX
XX KW Farnesyl transferase; inhibitor; cancer; ras; p21.
XX
XX OS Homo sapiens.
XX
XX PN US5420245-A.
XX
XX PD 30-MAY-1995.
XX
XX PF 18-APR-1990; 90US-0510706.
XX
XX PR 16-JAN-1992; 92US-0822011.
XX
XX PR 18-APR-1990; 90US-0510706.
XX
XX PR 20-NOV-1990; 90US-0615715.
XX
XX PR 03-APR-1992; 92US-0863169.
XX
XX PA (TEXA ) UNIV TEXAS.
XX
XX PI Brown MS, Goldstein JL, Reiss Y;
XX
XX DR WPI; 1995-206308/27.
XX N-PSDB; AAQ94412.
XX
XX PT New farnesyl transferase inhibitor peptide(s) - based on farnesyl
XX acceptor substrate carboxy terminal sequences, used for the
XX treatment of cancer
XX
XX PS Example 4; Column 55-58; 55pp; English.
XX
XX CC AAR77841 is the alpha subunit of human farnesyl transferase which
XX is involved in the farnesylation of various cellular proteins
XX including the cancer related ras proteins. It is used to produce the
XX complete farnesyl transferase molecule which is used to demonstrate
XX the effectiveness of peptide inhibitors capable of inhibiting
XX farnesyl transferases. The peptide inhibitors are useful for

```

CC treating cancers and ras-related cancers in particular.

XX Sequence 379 AA;

Query Match 87.2%; Score 34; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
|||||

Db 270 VPHNES 275

RESULT 14

AAW04431
ID AAW04431 standard; Protein; 379 AA.

XX AC AAW04431;

DT 30-JUL-1997 (first entry)

XX Human farnesyl transferase enzyme alpha subunit.

XX Farnesyl transferase; inhibitor; cancer; tumour; neoplasia; prenyl;
KW ras protein; K-ras B; malignant; detection; identification.

XX Homo sapiens.

XX WO9634113-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96WO-US05969.

XX 27-APR-1995; 95US-0429964.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Brown MS, Goldstein JL, James GL;

XX WPI; 1996-497642/49.

XX N-PSDB; AAT38710.

XX Assay for farnesyl transferase activity - by determining ability to
PT transfer farnesyl moiety to K-Ras B protein, partic. useful for
PT identifying inhibitors

XX Example 4; Page 151-154; 257pp; English.

XX AAW04431 shows the alpha subunit of a farnesyl transferase (FT) enzyme
CC derived from a human retinal cDNA library. The enzyme was used in a
CC method for identifying FT inhibitors. The method involved screening
CC candidate compounds for the ability to inhibit the transfer of a
CC farnesyl moiety to a K-ras B protein. FT inhibitors act by blocking
CC the attachment of prenyl groups to ras proteins in malignant cells of
CC patients suffering from cancer or precancerous states, and as such are
CC used to treat such conditions.

XX Sequence 379 AA;

Query Match 87.2%; Score 34; DB 17; Length 379;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6
|||||

Db 270 VPHNES 275

RESULT 15

ABB08436
ID ABB08436 standard; Protein; 379 AA.

XX

AC ABB08436;

XX 07-MAY-2002 (first entry)

XX Protein sequence 2 relative to the farnesyltransferase of the invention.

XX Farnesyltransferase; enzyme.

XX Unidentified.

XX KR98075770-A.

XX 16-NOV-1998.

XX 01-APR-1997; 97KR-0012067.

XX 01-APR-1997; 97KR-0012067.

XX (GLDS) LG CHEM LTD.

XX Moon GD, Kim MJ, Chung HH;

XX WPI; 2000-020309/02.

XX N-PSDB; ABA98898.

XX Farnesyltransferase having histidine tag and process for preparing the
PT same.

XX Disclosure; Page 16-17; 23pp; Korean.

XX The invention relates to a farnesyltransferase with a histidine tag,
CC and methods for preparing it. The current sequence represents protein
CC sequence 2 relative to the farnesyltransferase of the invention.

XX Sequence 379 AA;

Query Match 87.2%; Score 34; DB 21; Length 379;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPHNES 6

Db 270 VPHNES 275

Search completed: May 6, 2003, 15:01:04
Job time : 33.85 secs

THIS PAGE BLANK (uspto)